NOTE TO USERS

The original manuscript received by UMI contains indistinct, slanted and or light print. All efforts were made to acquire the highest quality manuscript from the author or school. Microfilmed as received.

This reproduction is the best copy available

UMI

UNIVERSITY OF ALBERTA

Auroral Current Systems Studied Using Automated Forward Modelling

BY

Martin Connors ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment
of the requirements for the degree of Doctor of Philosophy.

DEPARTMENT OF PHYSICS

Edmonton, Alberta

SPRING 1998

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-29027-1

Canadä

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled Auroral Current Systems Studied Using Automated Forward Modelling submitted by Martin Connors in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Gordon Rostoker

Douglas P. Hube

J. Anthony Kernahan

Douglas R. Schmitt

Edward P. Lozowski

Robert L. McPherron

January 30, 1998

Abstract

A new method, Automated Forward Modelling, has been developed for studying auroral zone currents using magnetic data. A three-dimensional forward model approximation is used for these currents. The forward model incorporates the magnetic effects of the global electrojets and one or more localized currents. Each current system is represented by a downward and an upward current sheet joined along an ionospheric path. A good reproduction of ground-based magnetograms is obtained by optimization of parameters, and a physical meaning may be attached to all parameters used. The ground measurements are not very sensitive to Region 1/2 field-aligned currents and their ionospheric closure. However, optimization including such currents has been successful in extracting meaningful parameters for them, primarily if data from the polar cap are included. One or more parameters may also be used to model large-scale fields of magnetospheric or atmospheric origin. In practice it is possible to reproduce magnetograms with a number of parameters approximately one-third of the number of data points: generally about 30 ground stations are used and the three magnetic components at each give rise to 90 variables, and the data set can be well represented by 31 parameters.

The technique has been applied with the intention of studying the time development and associations to other auroral phenomena of current systems associated with substorms. The CDAW 9B period of April 2/3 1986 to approximately 0300 UT has been analysed in great detail with multiple sources of data. Two CDAW (Coordinated Data Analysis

Workshop) substorms have been analysed in less detail: those of April 1 1986 and of May 3, 1986. A comprehensive data set is analysed including ground and spacecraft magnetometers and global auroral imaging. Good correspondence is obtained between current locations inferred from magnetograms and those inferred as areas of upward current on images. The time development of the global electrojets as far as current and location conforms to that generally known to be associated with substorms. The formation of intense current 'wedges' near magnetic midnight is reflected in the model parameters, with the currents quantitatively specified. Two smaller substorms have also been studied on November 8, 1993, and on August 8, 1991.

The modelling studies and analysis of ancillary data suggest that concentration of current near the poleward border of the auroral oval after substorm onsets is an important phenomenon, and generally support what has become known as the 'double auroral oval' hypothesis.

Preface

"It is a capital mistake to theorize before one has data" - Sherlock Holmes

Magnetospheric substorms remain incompletely understood. The general knowledge that energy injected into the magnetosphere from the solar wind is dissipated at least in part through the substorm phenomenon has not translated itself into a complete understanding of the physical processes involved. In that sense studies of auroral zone currents during substorms might be viewed as worthwhile in hopes of elucidating details of the overall process.

More specifically, however, such studies should clarify a controversial point which is very fundamental. It remains unclear to what degree the dissipation of solar wind energy has continuous and discrete components. Studies of magnetic field perturbations are a good tool for determining this due to their ability to reveal electric currents on a global scale. The refinement of techniques for extracting information from the magnetic measurements leads to use of these measurements to answer this specific question.

The overall view of substorms remains controversial after many years of work [Lopez, 1993; Lui 1991] with at least 5 models being championed. Much of the attempt to correlate phenomena observed from the ground (which are largely of ionospheric origin) with in-situ observations from spacecraft strategically placed (in the few 'lucky' cases) is based on secondary indicators of the ionospheric currents. These may be luminosity enhancements or magnetic perturbations, for example. The ability to obtain physical parameters of ionospheric currents for such correlative studies is of obvious value, making more direct comparison of data sets possible. Due to the use of 'indices' derived in a complex manner, one could argue that there are often in fact two steps between the physical parameters of current systems and their common parametrization. An example of this is seen in the cited Lopez paper where magnetic fields at the GOES 5 spacecraft in equatorial geosynchronous orbit are compared with the value of the AE index. That index in turn is claimed to represent primarily perturbations seen at Great Whale River (Poste de la Baleine); those perturbations in turn represent only one component of the magnetic field. Clearly a reliable solution for the ionospheric currents causing the perturbations would allow one to make a desired, more direct, connection between the data sets. The AE index is widely used, yet depends on physical processes in a complex fashion . It will be seen often in this study that behaviour of such auroral electrojet indices can be anomalous at times due to the changing dominance of a particular station or current system.

The basic nature of the substorm process has been characterized somewhat schematically in the past, in part due to the secondary inferences used. There are thought to be explosive characteristics associated with certain phases of the phenomenon and these in turn have led to consideration of physical processes involving magnetic fields which may lead to explosive release of energy, in particular magnetic reconnection [Akasofu, 1994]. Very sudden field changes seen at individual stations are indeed suggestive of an explosive process. It must be recalled, however, that the magnetic field due to a current depends very sensitively on geometric parameters. This makes the deconvolution of temporal and spatial changes difficult. A synthetic approach which automatically incorporates geometric parameters at many points has some chance of addressing this problem and allowing the types and role of reconnection to be better specified within the grand scheme of things.

Other aspects must be considered. The present time is one of transition in technique. Data storage and manipulation is increasingly done digitally, with numerous advantages. The speed of computers available for data analysis has increased so rapidly that what started as a project for the fastest machines on campus now was most effectively completed on a desktop system in a private home. New generations of spacecraft and better ground networks in some regions (worse in others, though) promise rich troves of data. The new techniques can be expected to be efficiently applied to large amounts of data and patterns revealed which may have escaped the past 'event studies' done with less quantitative techniques.

The present project has developed a new technique and shown it being used, with varying degrees of success, in modelling the causes of near-Earth magnetic perturbations. Hopefully the approach demonstrated will find a niche and be used in answering the larger questions which remain unanswered about the Sun-Earth interaction.

Acknowledgements

I acknowledge the contributions made by my supervisor Dr. Gordon Rostoker and by my examining committee in the completion of this work. Within the Department of Physics, trusted advice and supportive comments from Dr. Douglas Hube were particularly appreciated. Continued use of the computer network which I developed while working in the Space Science Laboratory was possible and fruitful due to the efforts of Sherwood Botsford. Some of the work was done on computers set up with assistance from Terry Tansky of Athabasca University.

Continued support from my wife Sherry and the pleasure of family life with my children Diane and Jade have made the sacrifices needed for success in studies appear in context as only a small part of life. Without the assistance of grandmothers Bee Connors and Diane Gene family demands would have overwhelmed. Family friend Mary Menard has also been very helpful with child care.

A valued support network existed in the form of scientific colleagues, many of whom I met through, and collaborated with, on the Internet. I am grateful to those who provided data as listed below, or reprints or preprints of their work, and in particular to Delores Knipp for sending a copy of her entire thesis. In addition to having helpfully provided data through the course of the work, Barbara Emery supplied custom AMIE data, on short notice at the end of the project, for use in comparison of methods.

The opportunity to teach has been another satisfying aspect of this undertaking. Supportive environments in the Minerva Institute for Seniors' Studies, the University of Alberta Faculty of Extension and Faculté Saint-Jean, and at Athabasca University have contributed to the rounding out of my development.

My graduate student colleagues and the support staff in the Space Science Laboratory have cheered me and generated a friendly atmosphere under pressure. I particularly note the contributions to my sense of balance made by Susan Skone and Eric Donovan.

In the midst of this work, I was saddened by the untimely passing of Dr. Barry Harrold. Having watched his development from young student to outstanding scientist, I regret that I will not have him as a colleague in research.

Finally, I am grateful for financial support in the undertaking of my doctorate, initially provided by the Natural Sciences and Engineering Research Council of Canada in the form of a scholarship. Additional support came from the University of Alberta in the form of financial and travel awards. Summer support was supplied through the NSERC funds of Dr. Gordon Rostoker, who also assisted with conference travel.

In a study based on analysis of data, it is important to have access to reliable data. It is a pleasure to acknowledge the following sources of data: even in cases where the data is in the public domain or was purchased, there were helpful people behind the streams of bits.

A list of those who have contributed follows, in no particular order: O. Rasmussen, E. Friis-Christensen, T. Moretto, and P. Hesselberg, Denmark; D. Milling and T. Clark, Great Britain; D. Herzog, USGS; G. Reeves, LANL; M. Peredo, ISTP project; F. Rich and J. Rodriguez, DMSP project; B. Emery, NCAR; S. Murphree, R. Elphinstone, and G. Enno, University of Calgary; L. Lanzerotti; A. Kotikov, Russia; T. Hughes and F. Creutzberg, CSA; Ari Viljanen, FMI (Finland); and R. Farlee for contribution of an IGRF routine.

CANOPUS data was supplied by the Canadian Space Agency (CSA). The CANOPUS instrument array was constructed and is maintained and operated by the Canadian Space Agency for the Canadian scientific community.

The length of this thesis, or any errors, are not attributable to any contributors, but my responsibility alone. I regret any omissions from the above list, which is large. There are many others to thank, and many kindnesses to hope to repay.

Table of Contents

1. Introduction 1
a. Solar Wind and Magnetosphere 1
b. Ionosphere 10
c. Magnetosphere-Ionosphere Coupling 11
d. Configuration of the Magnetosphere 14
e. Substorms 31
f. Electric Current Systems for Substorm Models 39

2. Physical Processes 41
a. Physical State of Near-Earth Matter 41
b. Magnetohydrodynamic Representation 41
c. Conductivity, Electric Field, Electric Current 42
d. Induction Effects 44
e. Substorm Time Scale 45

3. Data Availability and Preparation 51
a. Fields - Remote and In-Situ 51
    1. Magnetic Field 51
    a. Magnetometers 51
    b. Ground-Based Magnetometers 52
    c. Spacecraft Magnetometers 59
b. Particles 62
c. Imaging 64
    1. Viking 65
    2. Dynamics Explorer 66
    3. DMSP 66
    4. POLAR 66
    5. Meridian Scanning Photometry 67
d. Magnetic Data and Electric Currents 67
e. Coordinate Systems and Transformations 71

4. Modelling Techniques 79
a. Basic Problem 80
    1. Biot-Savart Integral 81
    2. Earth Induction 81
    3. Simplified Models 83
b. KRM 88
c. AMIE 91
d. Kisabeth's Modelling Approach 97
    1. Gaussian Integration 98

2. Components of Forward Model Current Systems 99
a. Toroidal Currents 101
b. Poloidal Currents 101
3. Application of Manual Forward Modelling 108

5. Optimization of Models 110
a. Parameters 111
b. Forward Model Choice 113
1. Parameters of East-West (Toroidal) Systems 115
2. Parameters of North-South (Poloidal) Systems 119
3. Constraints on Parameters 120
c. Optimization 123
d. Uniqueness 130
e. Error Bounds 132
f. Selection Criteria and Data Availability for Events 134
g. Optimized Fitting of Magnetic Data 135
h. Comparison with Other Data about the Events 151

6. CDAW 9B - April 2-3, 1986 152
a. Chronological Summary 152
b. Conditions preceding activity 158
c. Start of active period 159
d. First major onset 176
e. Small-scale behaviour in western wedge 208
f. Northerly onset 210
g. Second major onset 229
h. DMSP passes 249
i. Recovery 259
j. Later activity 260
k. Summary 265

7. CDAW 9A - April 1, 1986 266
a. Chronological Summary 266
b. Evening sector electrojet 270
c. Morning/midnight sector electrojets 277
d. Three-dimensional modelling 283
e. Summary 287

8. CDAW 9C - May 2-3, 1986 288
a. Background 290
b. Chronological Summary 302
c. Modelling Results 302

d. Summary 325

9. August 8, 1991 327
a. Data Preparation 328
b. Chronological Summary 330
c. Meridian chain results 334
d. Wedge modelling results 352
e. Global modelling results 357
f. Summary 364

10. GEM Event - November 8, 1993 366
a. Overview 366
b. Chronological Summary 368
c. Modelling results 371
d. Summary 387

11. Discussion and Conclusions 388
a. Characteristics of the Substorm Current Wedge 388
b. Aspects of the Global Electrojets near Onset Time 390
c. Other Aspects of Geomagnetic Phenomenology 393
d. Use of Activity Indices 393
e. Technical Considerations 395
f. Contextual Assessment of AFM 397
g. Suggestions for Further Work 410
h. Synopsis 410

References 411

List of Figures

1.1 Solar wind fluid properties observed with the IMP-8 satellite 2

1.2 Solar wind magnetic properties observed with the IMP-8 satellite 3

1.3 Thermal and magnetic pressure sum in the solar wind and the magnetotail 5

1.4 Dayside magnetopause crossings by GOES spacecraft on June 12 1990 6

1.5 Solar wind parameters on June 12 1990 7

1.6 Midlatitude H components on March 22 1979 8

1.7 Magnetic field at GEOS 2 synchronous spacecraft on March 22, 1979 9

1.8 Instantaneous view of the auroral oval from the Viking spacecraft 12

1.9 Borders of auroral oval current flow in one-hour averaged magnetic data 13

1.10 Borders of auroral oval current flow at high activity level 13

1.11 Canonical view of field-aligned currents from the Triad satellite 15

1.12 Potential patterns as projected onto the northern hemisphere 16

1.13 Electric fields as measured on polar passes of the DE-2 satellite 17

1.14 Magnetopause positions as projected onto the plane of Earth's orbit 20

1.15 Simultaneous observations of the two lobes from ISEE-1 and ISEE-2 23

1.16 ISEE 2 pass through the current sheet on April 2-3 1986 24

1.17 Particle data as ISEE 2 traverses the neutral sheet on April 2-3 1986 26

1.18 Cross-sectional pictorial synthesis of Earth's magnetosphere 31

1.19 Stage III of the auroral substorm expansion phase according to Akasofu 32

1.20 View of substorm current wedge as deduced by Clauer and McPherron 34

1.21 Northward magnetic perturbations at Meanook on November 6-7 (UT) 1997 35
1.22 Northward and earthward components at GOES 5 on April 3 and 4 1986 37
1.23 Electron fluxes in the 30 to 300 keV range at syncronous satellite 1984-037 38
1.24 Simplified ionospheric current systems for substorm models 40
2.1 Conductivities within and outside an auroral arc 43
2.2 Prototypical relation of polar magnetic substorms to geomagnetic bay 45
2.3 Comparison of IMF B_z and Auroral Electrojet indices 47
2.4 Behaviour of D_{st} Index during a period of storm activity 48
3.1 Magnetograms from the CANOPUS array for April 11, 1997 53
3.2 S_q ionospheric current system under equinoctial conditions 55
3.3 Averaged quiet day variations showing principally S_q variations 56
3.4 Perturbations observed at Fredericksburg 56
3.5 Comparison of two methods of correcting night time variations 57
3.6 Locations of the 12 standard auroral electrojet stations 58
3.7 Geographic X and Y components observed by the DMSP/F7 magnetometer 61
3.8 Dispersive injections of protons into the synchronous orbit region 63
3.9 Conjugacy of Husafell and Syowa Stations 68
3.10 Photographs of aurora simultaneously taken at conjugate points 70
3.11 Superposed northern and southern hemisphere maps showing conjugate stations 71
3.12 Curves of CGM and centred dipole latitudes as a function of geodetic longitude 74
3.13 Coordinates of DMSP F7 satellite footprints on a northern hemisphere pass 76

3.14 Latitude of DMSP F7 spacecraft on April 2 1986 77

4.1 Averaged ground equivalent current and TRIAD magnetic disturbance vectors 86

4.2 Legendre polynomial $P_{20}(x)$ for $x=\cos\theta$ in the interval $\theta = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ 95

4.3 Legendre polynomial $P_{20}(x)$ for x in the interval [-1,1] 99

4.4 Magnetic perturbation at 70° N due to westward current flow in a 180° system 100

4.5 Latitude profile for ionospheric current in a north-south system 102

4.6 Jump in Y_m component for long 1 MA downward sheet current at 75° N 103

4.7 Jump in Y_m component for long 1 MA upward sheet current at 65° N 103

4.8 Total perturbation associated with field-aligned sheet currents of previous figures 104

4.9 Latitude profile of total perturbation for north-south current flow of 1 MA 104

4.10 Ground effects of two 160° long north-south current systems 105

4.11 Ground effects of two 160° long NS current systems from Kisabeth [1979] 106

4.12 Ground perturbations associated with a long north-south current 106

4.13 Ionospheric current inferred from radar compared to ground perturbation 107

4.14 Superposed epoch results of Hughes [1978] for the noon sector 109

5.1 Single east-west current system illustrating its parameters 116

5.2 Reversal of a single bounding current sheet to result in a 'bowtie' 117

5.3 Optimal fit of a curved current system by a 'bowtie' system 118

5.4 North-south current system as specified by seven parameters 119

5.5 Fukushima equivalent currents for a field-aligned current pair 131

5.6 Perturbations off the end of a long electrojet, modelled by overhead current 134

5.7 Behaviour of χ^2 as a function of station number 137
5.8 Chi-squared on linear scale throughout test convergence 138
5.9 Logarithmic graph of χ^2 in test convergence 139
5.10 Value of scale parameter λ during convergence of test case 140
5.11 Convergence of total current in system from poor to good value 141
5.12 Uneven convergence of system width parameters toward target values 142
5.13 Variation of latitudes of centre points during convergence 142
5.14 Variation of longitudes of extremal points of current sheets during iterations 143
5.15 Count of convergence failures in a full data run of the optimization 144
5.16 Scale length parameter l in an actual program run 145
5.17 Behaviour of χ^2 through a real optimization run 146
5.18 χ^2 throughout convergence of a realistic but 'poor' initial model 147
5.19 Initial current configuration for realistic global test model 149
5.20 Final solution, showing perturbation magnetic field, for AFM global run 150
6.1 IMP 8 solar wind data for April 2 1986 157
6.2 Viking view of expansive phase activity on April 2 1986 157
6.3 X component at Barrow from 10 to 16 UT April 2 1986 159
6.4 Viking images from orbit 219, April 2 1986 160
6.5 IMP 8 GSM field on April 2-3 1986 (X and Y components) 161
6.6 IMP 8 GSM field on April 2-3 1986 (Z component and total field) 161
6.7 Auroral Electrojet Indices on April 2-3 1986 162

6.8 X Component Perturbations at Greenland stations on April 2 1986 163
6.9 X Component Perturbations at Kap Tobin on April 2 1986 163
6.10 XYZ Perturbations at Isafjordur, Iceland, April 2-3 1986 164
6.11 Viking images in the vicinity of pre-onset Greenland by Camera 0 (LBH) 165
6.12 Contrast-enhanced Viking image showing tilted midnight sector electrojet 166
6.13 XYZ magnetogram for Kap Tobin on April 2/3 1986 167
6.14 $X_M Z$ magnetogram for Iqaluit on April 2 1986 168
6.15 XYZ magnetogram for Soroya on April 2 1986 169
6.16 XYZ magnetogram for Saint John's for April 2/3 1986 169
6.17 IMP-8 solar wind parameters on April 2-3 1986 171
6.18 IMP-8 bulk flow speed on April 2-3 1986 171
6.19 Syowa Pi 2 pulsations on April 2/3 1986 173
6.20 Syowa Pi 1/Pc 2 pulsations on April 2/3 1986 173
6.21 Electron flux at the SC 1984-037A synchronous satellite 174
6.22 Proton flux at the SC 1982-019A synchronous satellite 175
6.23 Perturbations at Chambon-La-Forêt and St. John's 176
6.24 X_M component perturbations at Narssarssuaq and at Husafell, Iceland 177
6.25 Stations used in modelling for 0100-0200 UT April 3 1986 181
6.26 Initial conditions used in modelling for 0100-0200 UT April 3 1986 182
6.27 Magnetic perturbations at Halley Bay on April 2-3 1986 184
6.28 Results of modelling from 0104 to 0126 UT April 3 1986 185

6.29 Results of modelling at 0100 UT April 3 1986 186
6.30 Results of modelling at 0105 UT April 3 1986 187
6.31 Results of modelling at 0110 UT April 3 1986 188
6.32 Results of modelling at 0115 UT April 3 1986 189
6.33 Results of modelling at 0120 UT April 3 1986 190
6.34 Results of modelling at 0125 UT April 3 1986 191
6.35 Viking images from 0104 to 0106 UT April 3 1986 195
6.36 Viking images from 0107 to 0109 UT April 3 1986 196
6.37 Viking images from 0110 to 0111:26 UT April 3 1986 196
6.38 Viking images from 0112 to 0114 UT April 3 1986 197
6.39 Viking images from 0115 to 0117 UT April 3 1986 198
6.40 Viking images from 0117:28 to 0119:28 UT April 3 1986 198
6.41 Viking images from 0120 to 0122 UT April 3 1986 199
6.42 Viking images from 0123 to 0125 UT April 3 1986 199
6.43 Magnetic perturbations at Poste-de-la-Baleine from 0 to 3 UT on April 3 1986 200
6.44 Modelling results for electrojet borders of Russian chain on April 3 1986 202
6.45 Latitude profile of Russian chain at 0 UT April 3 1986 203
6.46 Latitude profile of Russian chain at 0050 UT April 3 1986 204
6.47 Latitude profile of Russian chain at 0105 UT April 3 1986 205
6.48 Current across Russian chain at 0050 UT 206
6.49 Current across Russian chain at 0050 UT April 3 1986 207

6.50 χ^2 for two fitting runs to Russian chain 208

6.51 Viking images of 'wedgelet' from 0116 to 0118 UT April 3 1986 209

6.52 Results of modelling at 0130 UT April 3 1986 211

6.53 Results of modelling at 0135 UT April 3 1986 212

6.54 Results of modelling at 0140 UT April 3 1986 213

6.55 Results of modelling at 0145 UT April 3 1986 214

6.56 Results of modelling at 0150 UT April 3 1986 215

6.57 Results of modelling at 0155 UT April 3 1986 216

6.58 Results of modelling at 0200 UT April 3 1986 217

6.59 Perturbations at Bjørnøya on April 3 1986 219

6.60 Perturbations at Heiss Island on April 3 1986 220

6.61 Viking images of northerly onset in the period 0138 to 0140 UT April 3 1986 .. 221

6.62 Viking images of northerly onset in the period 0141 to 0143 UT April 3 1986 .. 221

6.63 Viking images of northerly onset in the period 0143 to 0145 UT April 3 1986 .. 222

6.64 Position of AMPTE/IRM satellite on April 2-3 1986 223

6.65 Elevation of the magnetic field at AMPTE/IRM on April 3 1986 223

6.66 GSE azumith of the magnetic field at AMPTE/IRM on April 3 1986 224

6.67 Magnitude of the magnetic field at AMPTE/IRM on April 3 1986 224

6.68 Proton bulk flow speed at AMPTE/IRM on April 3 1986 225

6.69 Proton temperature at AMPTE/IRM on April 3 1986 225

6.70 Electron flux at satellite SC 1984-037A on April 3 1986 227

6.71 Perturbations at West Greenland station Frederikshåb, 72° invariant latitude 228

6.72 Viking images showing sudden brightening in West Greenland 229

6.73 Y_M Perturbations at Glenlea, Fredericksburg, and Fresno on April 3 1986 230

6.74 Initial conditions for modelling during the hour 0200 to 0300 UT April 3 1986 231

6.75 Modelling results at 0200 UT April 3 1986 232

6.76 Modelling results at 0205 UT April 3 1986 233

6.77 Modelling results at 0210 UT April 3 1986 234

6.78 Modelling results at 0215 UT April 3 1986 235

6.79 Modelling results at 0220 UT April 3 1986 236

6.80 Modelling results at 0220 UT April 3 1986 including Region 1/2 currents 237

6.81 Modelling results at 0230 UT April 3 1986 238

6.82 Modelling results at 0235 UT April 3 1986 239

6.83 Modelling results at 0240 UT April 3 1986 240

6.84 Modelling results at 0245 UT April 3 1986 241

6.85 Modelling results at 0250 UT April 3 1986 242

6.86 Modelling results at 0255 UT April 3 1986 243

6.87 Perturbations at high latitude station Iqaluit from 0 to 3 UT on April 3 1986 247

6.88 DMSP Y component as a function of latitude on April 2 1986 (Pass 24) 250

6.89 Latitude profile at EISCAT meridian for 20 UT April 2 1986 251

6.90 DMSP Y component as a function of latitude on April 2 1986 (Pass 26) 252

6.91 Latitude profile at EISCAT meridian for 21.6 UT April 2 1986 253

6.92 DMSP Y component as a function of latitude on April 2 1986 (Pass 28) 254
6.93 Latitude profile at EISCAT meridian for 23.4 UT April 2 1986 254
6.94 DMSP Y component for Pass 2 April 3 and Pass 1 April 4, 1986 255
6.95 DMSP Y component for Pass 4 April 3 and Pass 3 April 4, 1986 257
6.96 DMSP Y component for Pass 3 April 3 and Pass 2 April 4, 1986 258
6.97 DMSP Y component for Pass 6 April 3 and Pass 5 April 4, 1986 259
6.98 Viking auroral image (Camera 0 LBH emission) at 0607 UT April 3 1986 260
6.99 Viking camera 1 image immediately before expansive phase onset 261
6.100 Viking camera 1 image immediately after expansive phase onset 261
6.101 AL index and magnetic perturbations 0600-0800 April 3 1986 262
6.102 AL index and magnetic perturbations at Gillam April 3 1986 263
6.103 AL index and magnetic perturbations at PDB and NAQ, April 3 1986 264
7.1 IMP-8 GSM magnetic fields on April 1 1986 268
7.2 Auroral electrojet indices on April 1 1986 269
7.3 Gain 100 Viking images of Orbit 213, April 1 1986 270
7.4 Westward current over Greenland from 18 UT to 1930 UT April 1 1986 271
7.5 Electrojet borders over Greenland from 18 UT to 1930 UT April 1 1986 272
7.6 Latitude profiles over Greenland at 1800 UT April 1 1986 273
7.7 Latitude profiles over Greenland at 1830 UT April 1 1986 273
7.8 Latitude profiles over Greenland at 1900 UT April 1 1986 274
7.9 Latitude profiles over Greenland at 1930 UT April 1 1986 274

7.10 Observed and modelled perturbations at Frederikshåb on April 1, 1986 275
7.11 Observed and modelled perturbations at Godthåb on April 1, 1986 276
7.12 Current over Russian chain from 18.5 to 19.5 UT April 1 1986 277
7.13 Borders of electrojet over Russian chain, 18.5 to 19.5 UT April 1 1986 278
7.14 Russian chain latitude profile at 18.5 UT April 1 1986 279
7.15 Russian chain latitude profile near onset time at 1850 UT April 1 1986 279
7.16 Russian chain latitude profile at 1900 UT April 1, 1986 280
7.17 Russian chain latitude profile at 1930 UT April 1, 1986 281
7.18 DMSP F7 Y component on April 1 1986 282
7.19 Global modelling results using eastward electrojet and a current wedge 284
7.20 Global modelling results using two electrojets and current wedge 286
8.1 Auroral electrojet indices from 18 UT May 2 to 12 UT May 3 1986 288
8.2 Conditions prior to major onset on May 3 1986 289
8.3 Vignetting-corrected Camera 0 Viking images on May 3 1986 291
8.4 XYZ magnetic perturbations at Poste de la Baleine on May 3 1986 293
8.5 Perturbations in local magnetic system at Siple, Antarctica, on May 3 1986 293
8.6 Local magnetic northward and Z components at Syowa and Husafell 295
8.7 Electron flux at satellite SC 1986-019 on May 3 1986 295
8.8 Centred dipole magnetic perturbations at Muonio on May 3 1986 296
8.9 Simulated data from the Halley Bay conjugate point on May 3 1986 297
8.10 Z at Halley Bay's conjugate point, STJ and NAQ on May 3 1986 298

8.11 Dynamics Explorer southern hemisphere images on May 3 1986 299

8.12 Magnetic field at GOES 5 and GOES 6 on May 3 1986 300

8.13 Dst index values for early May 1986 303

8.14 Ps 6 pulsations in the EISCAT meridian on May 3, 1986 305

8.15 Electrojet boundaries in the EISCAT meridian on May 3 1986 306

8.16 Eastward current over EISCAT Meridian on May 3 1986 308

8.17 Magnetogram for Narssarssuaq spanning May 2 to May 3 1986 309

8.18 Observed and modelled perturbations at Narssarssuaq on May 2 1986 310

8.19 Electrojet boundaries in Greenland on May 2 1986 311

8.20 Current over Greenland on May 2 1986 312

8.21 Latitude profile of H and Z in western Greenland at 2336 UT May 2 1986 313

8.22 Latitude profile of H and Z in western Greenland at 2345 UT May 2 1986 313

8.23 Perturbations and model for western Greenland at 2.6 UT on May 3 1986 315

8.24 Locations of stations used in modelling on May 2-3 1986 317

8.25 Global modelling results at 0105 UT May 3 1986 318

8.26 Global modelling results at 0112 UT May 3 1986 320

8.27 Modelling results including NS currents at 0112 UT May 3 1986 321

8.28 Global modelling results at 0122 UT May 3 1986 324

9.1 Midlatitude Y components on August 8, 1991 327

9.2 Perturbations at GOES-6 on August 8, 1991 332

9.3 Perturbations at GOES-7 on August 8, 1991 332

9.4 Meridian scanning photometry at CANOPUS stations on August 8 1991 333

9.5 Current over Churchill chain from 6.8 to 7.8 UT August 8 1991 335

9.6 Borders of electrojet over Churchill chain, 6.8 to 7.8 UT August 8 1991 335

9.7 Current over Churchill chain from 6.8 to 7.8 UT August 8 1991 336

9.8 Borders of electrojet over Churchill chain, 6.8 to 7.8 UT August 8 1991 337

9.9 Pi2 band pulsations at Pinawa and Island Lake on August 8 1991 339

9.10 Pi1 band pulsations at Pinawa and Island Lake on August 8 1991 339

9.11 Data and reconstructed magnetogram for Island Lake on 8 August 1991 340

9.12 Data and reconstructed magnetogram for Fort Churchill on 8 August 1991 341

9.13 Relationship between observed and modelled X data on 8 August 1991 342

9.14 Value of chi-squared in two modelling runs of the Churchill meridian 343

9.15 Electrojet boundaries in a two-electrojet model for August 8 1991 345

9.16 Current in each electrojet of the two-electrojet model for August 8 1991 345

9.17 Current through the Alaska meridian on August 8 1991 346

9.18 Electrojet borders in the Alaska sector on August 8 1991 347

9.19 Latitude profile of Alaska chain stations at 7.10 UT August 10 1991 349

9.20 Latitude profile of Alaska chain stations at 7.10 UT August 10 1991 349

9.21 Observed and modelled perturbations at College on August 8 1991 350

9.22 Observed and modelled perturbations at Talkeetna on August 8 1991 350

9.23 Station key for modelling for August 8 1991 351

9.24 Model of pre-onset conditions in North America on August 8 1991 354

9.25 Modelling results near time of onset at 0705.5 UT on August 8 1991 ... 355
9.26 Modelling results at 0710.5 UT on August 8 1991, after time of onset ... 356
9.27 Global modelling results at 0700.5 UT August 8, 1991 ... 359
9.28 Global modelling results at 0705.5 UT August 8, 1991 ... 361
9.29 Global modelling results at 0710.5 UT August 8, 1991 ... 362
9.30 Current through Churchill meridian from meridian and global modelling ... 364
10.1 AL and AU indices on November 2 to 11, 1993 ... 366
10.2 Magnetic observations at College and at Dawson on November 8 1993 ... 369
10.3 Dispersionless injection observed by SC 1989-097 on November 8, 1993 ... 369
10.4 IMP-8 IMF in GSM coordinates from 0600 to 1200 UT November 8, 1993 ... 370
10.5 Positions of GEM stations used for modelling for November 8, 1993 ... 373
10.6 Initial conditions for modelling for November 8 1993 ... 374
10.7 Modelling results at 1030 UT November 8 1993 ... 375
10.8 Five minute averaged data from Barrow on November 8, 1993 ... 376
10.9 One minute data from Barrow on November 8, 1993 ... 377
10.10 Modelling results at 1035 UT November 8 1993 ... 379
10.11 Modelling results at 1050 UT November 8 1993 ... 380
10.12 Modelling results at 1055 UT November 8 1993 ... 381
10.13 Modelling results including eastward electrojet at 1050 UT November 8 1993 ... 383
10.14 Modelling results including eastward electrojet at 1055 UT November 8 1993 ... 385
10.15 Total current in the eastward electrojet on November 8, 1993 ... 386

11.1 Positions of stations for AMIE modelling in Apex Magnetic Coordinates 398

11.2 AMIE output showing Hall current on April 3, 1986, at 0105 UT 399

11.3 AMIE output showing Hall current on April 1, 1986, at 1905 UT 402

11.4 Perturbations as observed and from AMIE for April 1, 1986 at 1905 UT 404

11.5 FAC associated with Hall current on April 1, 1986, at 1905 UT 406

11.6 FAC associated with Pedersen current on April 1, 1986, at 1905 UT 407

11.7 Observed and modelled X values from AFM and AMIE 408

List of Tables

1.1 Near Tail Plasma Parameters 25
5.1 Current Parameters for Two Current Modelling Run 122
5.2 Rules for Variation and Constraint of Parameters for Two Currents 122
5.3 Parameters of test (target) system and initial guess 138
6.1 Conditions preceding activity (Section 6.b) 153
6.2 Start of active period (Section 6.c) 154
6.3 First major onset (Section 6.d) 155
6.4 Northerly onset (Section 6.f) 155
6.5 Second major onset (Section 6.g) 156
6.6 Recovery (Section 6.I) 156
6.7 Later activity (Section 6.j) 156
6.8 Stations used in modelling for CDAW Event 9B 178
6.9 Stations used in making Russian meridian profiles 201
7.1 CDAW 9A Onset on April 1 1986 267
8.1 CDAW 9C Onset on May 3 1986 302
9.1 Parameters of Alaskan Stations 329
9.2 7 UT Onset on August 8 1991 331
10.1 1050 UT Onset on November 8 1993 371
11.1 Comparison of current across meridians in AMIE and AFM on April 400
11.2 Comparison of current across meridians in AMIE and AFM on April 403
11.3 Assessment of AMIE and AFM with regression parameters 409

Chapter 1. Introduction

This chapter sets the stage for what follows through a brief presentation of the known characteristics of near-Earth space, with an emphasis on the observational basis of our knowledge. Data are presented which lead to a generalized model of the interaction of the solar wind with the Earth. This model forms the framework for later discussions.

a. Solar Wind and Magnetosphere

The driving force behind the solar-terrestrial interaction is the solar wind. This flow arises from the thermal imbalance of the hot (temperatures exceed 10^6 K) corona of the Sun. This imbalance drives an outflow of plasma from near-solar regions which passes Earth's orbit and has now been observed by the Voyager spacecraft to continue past the outermost known planets. Coronal heating is likely due to the presence of magnetic fields in the solar plasma [Priest, 1995], and these also lead to the possibility of interaction with near-Earth plasmas and magnetic fields as will be discussed below. Figure 1.1 [National Space Science Data Center, 1989] illustrates the fluid properties of the solar wind as seen near Earth for the period March 27 to April 22, 1986. CDAW events A and B, which will be discussed in Chapters 7 and 6, respectively, occur on April 1 and April 3. The solar wind bulk speed, as shown, varies between approximately 300 and 700 km/s and a typical value would be 450 km/s. The density of ions (primarily protons, H^+) varies between 3×10^6 and 3×10^7 m^{-3} (shown as 3 to 30 per cc on the figure) and a value of 10^7 m^{-3} could be taken as representative. Much higher values are possible, as suggested by occasional observations of 10^8 or more particles per m^3 (see Chapter 10). The temperature varies from 10^4 K to about 3×10^5 K, again typical values but considerably less than in the solar coronal source regions. A temperature of 10^5 K corresponds to a particle thermal energy of about 10 eV (one eV being 1.6×10^{-19} J): both this and the flow speed of order 0.001 c (c being 3×10^8 m/s) mean that the solar wind is a nonrelativistic flow. The kinetic energy associated with ion bulk flow is about 1 keV per particle, so dynamic pressure greatly exceeds thermal pressure in this flow. As might then be expected, the sound speed, $c_s = \sqrt{\frac{\gamma p}{\rho}}$, with $\gamma = \frac{5}{3}$ the ratio of specific heats at constant pressure, p the thermal pressure, and ρ the mass density, is less than the flow speed. Since the plasma must be approximately neutral and is primarily made of hydrogen, the number density of electrons is approximately equal to that of protons. Thus, the total number density is about twice that, n, of protons, and the pressure is $p=2nkT$, k is the Boltzmann constant of 1.38×10^{-23} J K^{-1}, and T the observed temperature. This results in a thermal pressure of 3×10^{-11} Pa. The mass density is simply the number density times mass for protons since protons and electrons are the main constituents, with the mass of electrons negligible. Thus the mass density is about 2×10^{-20} kg/m^3, and the sound speed may be calculated as 50 km/s. As expected this is considerably less than the bulk flow speed. The solar wind is thus a supersonic flow. The actual bulk flow direction (not obtainable from the data shown) is approximately radial [Hundhausen, 1995]. The energy flux in the bulk flow is then $\frac{1}{2}\rho v^3$, which is readily calculated as 9×10^{-4} W/m^2, roughly 10^{-6} of the solar constant of 1 kW/m^2 which is mostly in the form of photons. Clearly the effects of the solar wind bulk flow will be minimal in areas able to interact with photons (such as those near the Earth's surface). This is not the case in ionized outer regions still primarily under the Earth's influence through its magnetic field: these regions are profoundly affected by the solar wind. However, the subtlety of solar wind effects as detected at the surface of the planet is primarily due to the small amount of energy carried in that flow compared to that of the photon flux.

Figure 1.1 Solar wind fluid properties observed with the IMP-8 satellite.

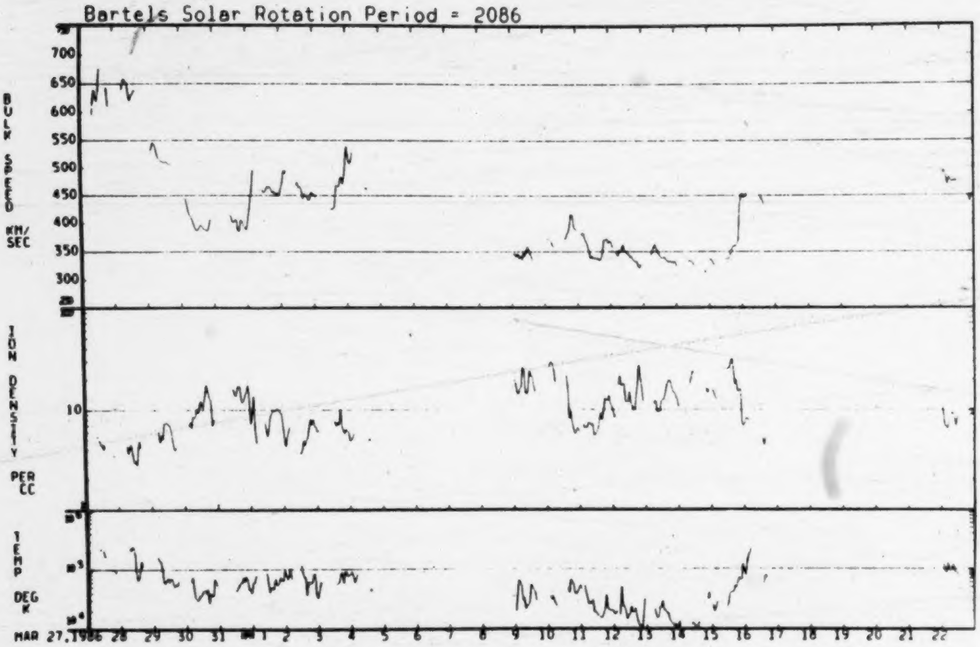
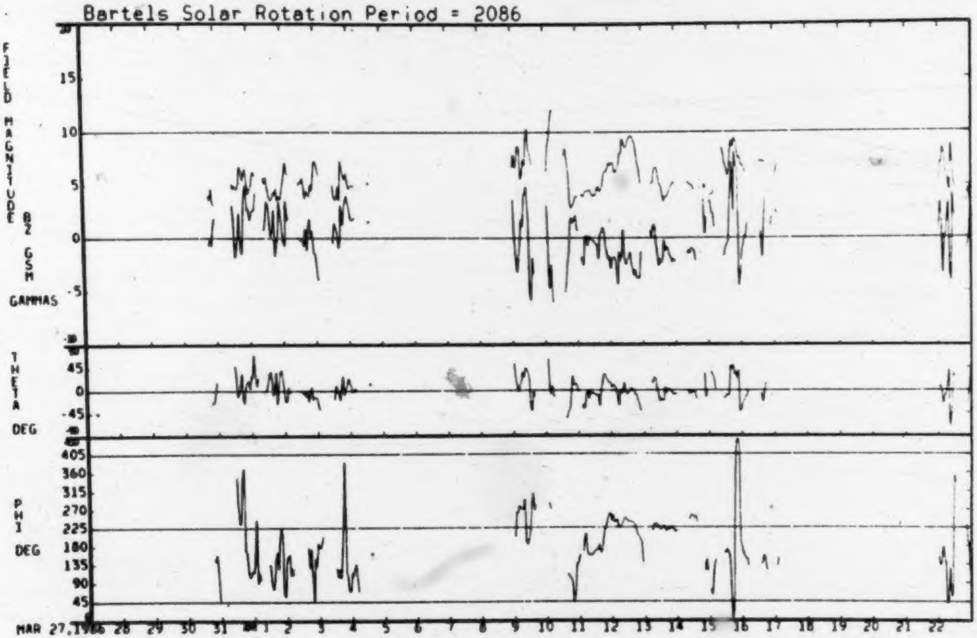


Figure 1.2 shows the observed magnetic properties of the solar wind during the period for which Figure 1.1 showed the fluid properties. The magnetic field associated with the solar wind is usually referred to as the interplanetary magnetic field (IMF) [Parks, 1991, p. 3]. The field magnitude near Earth is variable (here between about 3 and 12 nT) with 5 nT being a 'canonical' value. The indicated GSM Z component of the field has a mean which is slightly above zero in the period before 5 April and slightly below zero in the period 8-16 April. It is unsteady and crosses the zero line frequently. The GSM Z coordinate axis is parallel to the projection of Earth's dipole in a plane perpendicular to the ecliptic (see section 3.e) and the GSM Z component of the solar

wind field has been shown both theoretically and by statistical studies to be relevant to the question of energy coupling between the solar wind and the magnetosphere. In this figure, theta is the angle above the ecliptic of the field vector and phi is the angle of the ecliptic projection of the field from the Earth-Sun line. Careful inspection will show that the theta polar angle of the field vector shown does not correspond precisely to the arctangent of GSM Z: this angle is in fact the arctangent of GSE Z (again see section 3.e) and corresponds to the angle of the field to the ecliptic. The mean of theta seems to be near zero. The phi angle shown is that between the Earth-Sun line and the ecliptic projection of the field. Although subject to occasional large excursions, this angle is typically 135°. Although bulk flow is radial, the rotation of the Sun and the frozen-in field condition (see section 2b) cause magnetic field lines to be skewed at this angle from the Earth-Sun line. The magnetic field exerts a pressure and with a typical value of 5 nT, this pressure, $\frac{B^2}{2\mu_0}$, is 1.0×10^{-11} Pa, which is less than, although of the same order as, the thermal pressure. Both are small compared to the dynamic pressure or momentum flux of the bulk flow (which is $\rho v^2=4\times10^{-9}$ Pa)[1], as might be expected from the above discussion of energy content.

Figure 1.2 Solar wind magnetic properties observed with the IMP-8 satellite. Uppermost curve is field magnitude in nT (marked as gammas).



---

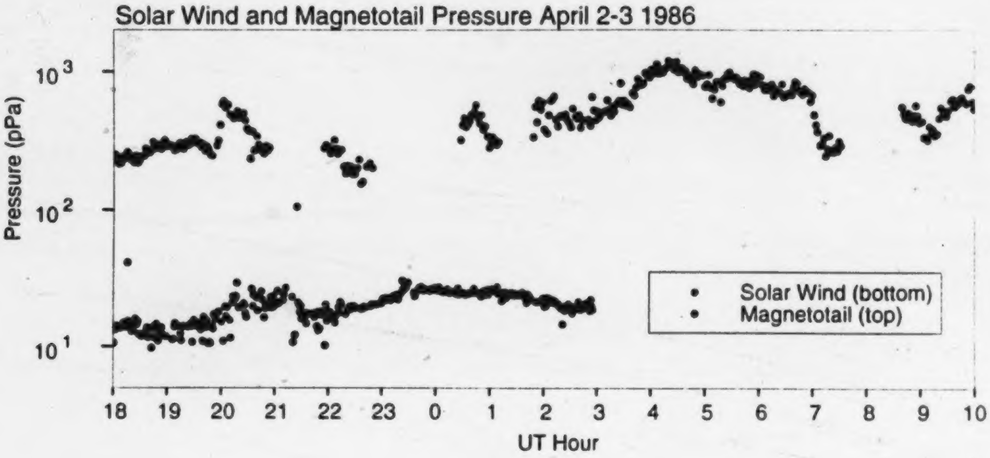
[1] Hundhausen [1995] cites 2.6×10^{-9} Pa. In SI units, 10^{-9} Pa=1 nPa.

The formation of a magnetosphere around the Earth may be understood by considering that the planet has an intrinsic magnetic field. The magnetosphere is a region whose existence is due to the tendency of this planetary magnetic field to exclude the solar wind. To a first approximation the Earth's intrinsic field is that of a dipole of magnetic moment 8×10^{15} Tm^3 or equivalently 30400 nTR$_e^3$ where R$_e$ is one Earth radius (6371 km). In the magnetic equatorial plane such a field is entirely vertical (perpendicular to the plane) and its strength is Mr^{-3}, where M is the dipole moment and r the radial coordinate. The magnetic pressure, $\frac{B^2}{2\mu_0}$, is thus $\frac{M^2r^{-6}}{2\mu_0}$. Earth's dipole is inclined at only 11° from the spin axis. This in turn is inclined at the 23.5° obliquity from the perpendicular to the ecliptic, so the subsolar point always lies relatively close to the magnetic equator. The maximum magnetic latitude of the subsolar point is observed at solstices and is then roughly 35°. A rough idea of the size of the magnetosphere can be obtained by balancing the two dominant sources of pressure above the subsolar point (i.e. along the Earth-Sun line) somewhere between the Earth and the Sun. The dynamic pressure dominated region of the solar wind and the internal field dominated dayside magnetosphere could be thought in this first order picture to exist in static equilibrium. Equating the magnetic pressure to the dynamic pressure at the subsolar point, the pressure balance would be such that $3.7\times10^{-4} = 4\times10^{-9}\ L_s^6$, where L_s is the 'standoff distance' expressed in units of Earth radii. This rough estimate[2] of standoff distance yields 6.7 R$_e$. In reality the dayside of the magnetosphere is compressed by the solar wind, enhancing the magnetic field over dipolar values, and thus the observed standoff distance is normally larger, about 10 R$_e$. It also follows that higher speed solar wind flows, with enhanced dynamic pressure, reduce the standoff distance. On the nightside the effect of the solar wind, rather than compressing the intrinsic field, is to sweep back plasma and associated magnetic flux into a long 'comet-like' magnetotail. That the dynamic pressure is the dominant effect in magnetotail formation is illustrated by Figure 1.3 which shows, using data from IMP 8 in the solar wind and ISEE in the magnetotail, that the static (thermal plus magnetic) pressure in the solar wind is one to two orders of magnitude less than that in the magnetotail. The much greater microscopic pressure in the magnetotail shows that solar wind dynamic pressure is needed to explain the confinement of the magnetotail. The component of solar wind dynamic pressure on the flanks of the magnetotail is almost entirely responsible for its confinement.

---

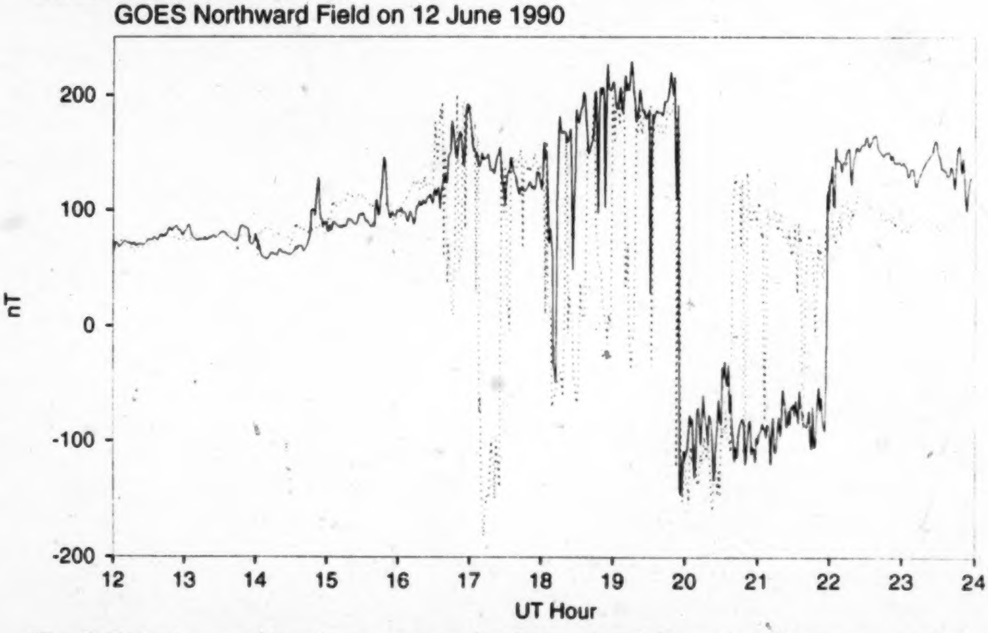
[2] The standoff distance varies as only the one-sixth power of dynamic pressure. With 2.6×10^{-9} Pa dynamic pressure, the standoff distance in this calculation changes to 7.4 R$_e$.

Figure 1.3 Thermal and magnetic pressure sum (in SI pPa units) in the solar wind and the magnetotail. Data gaps exist throughout the magnetotail data and in solar wind data starting at 3 UT.



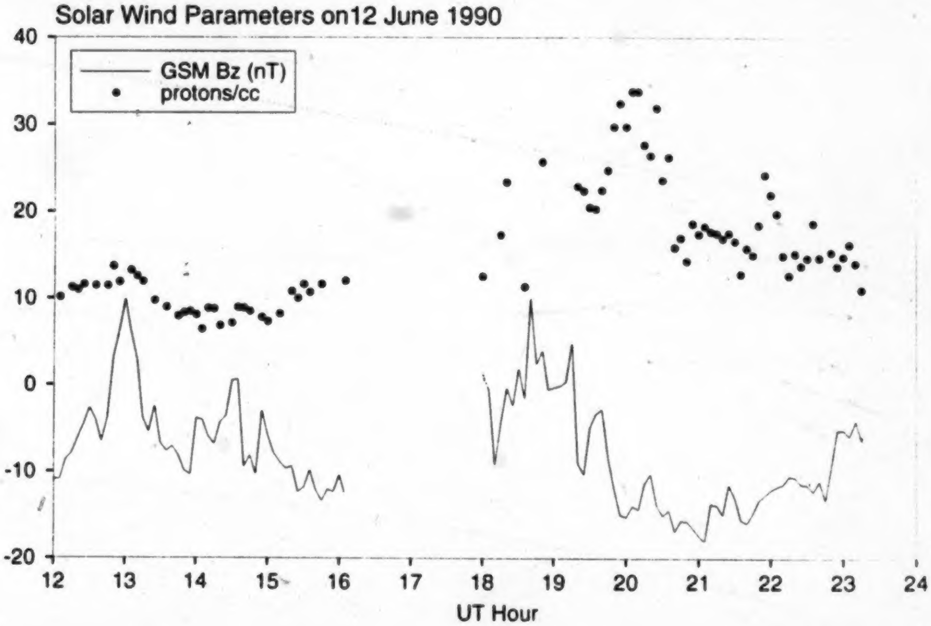
In reality the interaction between the magnetic field dominated magnetosphere and the dynamic pressure dominated solar wind is more complex than the simple balance at an imagined interface described above. The diversion of the supersonic solar wind flow occurs at a bow shock inside which conditions abruptly change from those of the solar wind. The bow shock is observed to be in front of the actual boundary of the magnetosphere, which is the region dominated by fields directly associated with the Earth. This boundary is called the magnetopause and the transition between internal and external fields is associated with a magnetopause current, usually referred to as a Chapman-Ferraro current. Spacecraft traversals of this current may be seen as abrupt magnetic field changes in Figure 1.4, which shows a case on June 12 1990 when the magnetopause had been pushed back to geosynchronous orbit. The abruptness of the transitions indicates that this current is very thin: in fact it is typically only a few hundred km thick [Hughes, 1995]. The field is seen to basically reverse on transition, rather than dropping to the very small interplanetary field value: an exterior region exists having fields roughly comparable in magnitude to those inside the magnetopause. This region between the magnetopause and the bow shock is referred to as the magnetosheath. On the dayside it can be thought of as containing slowed-down solar wind plasma and a magnetic field which is related to that of the solar wind. After 16 UT two GOES spacecraft (GOES 6 and 7) both detected fields stronger than the nominal 100 nT field at the equatorial plane in synchronous orbit, presumably evidence of compression of the magnetosphere. That compression occurred is further evidenced by the observation of magnetopause crossings at synchronous orbit, that is, only 6.6 R_e from the Earth as opposed to a more typical 10 R_e. On this date, mean solar noon for GOES 7 occurred at 18.84 UT, and for GOES 6 at 21.00 UT. Prior to about 20 UT, magnetopause crossings occur mainly at GOES 7, which was closer to the subsolar point. Field values during these crossings attain about -50 nT, except during one crossing by GOES 7 at 17.0 UT. By 20.5 UT, this spacecraft had rotated away from noon sufficiently to not continuously observe the magnetosheath, while GOES 6 remained immersed in it. Before further discussion, it is instructive to examine simultaneous solar wind data.

Figure 1.4 Dayside magnetopause crossings by GOES spacecraft on June 12 1990. Line representing northward component of magnetic field at GOES 6 is solid, at GOES 7 dotted.



The IMF had a southward component during much of the period shown, as seen in Figure 1.5. The proton density in the first part of the interval shown is a near-canonical 10^7 m^{-3} (shown as 10/cc), with stepwise increases to roughly three times this value by just after 20 UT. These data were obtained by the IMP-8 spacecraft which was in the solar wind at roughly 8 R_e sunward, 25 R_e duskward, and 20 R_e southward of the Earth (at 18 UT). The solar wind speed increased over several hours from a steady 550 km/s to nearly 800 km/s just after 18 UT and declined slowly thereafter (data not shown). Unfortunately the data gap seen in Figure 1.5, between roughly 16 and 18 UT, also is present in the speed measurements and does not permit details of the rise in solar wind speed to be specified.

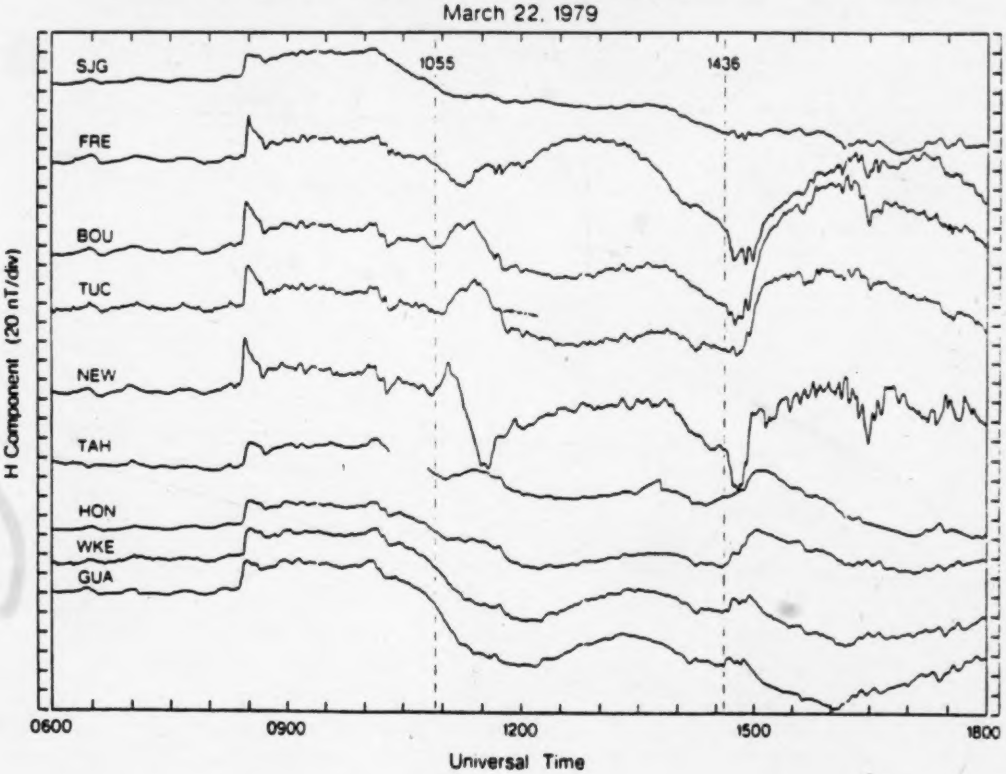
Figure 1.5 Solar wind parameters on June 12 1990. The northward component of the IMF in nT is shown as a solid line, proton number density per cm^3 as dots with a common numeric scale.



Taken together, the solar wind data and magnetopause crossings illustrate the three factors in the solar wind which may perturb the magnetosphere. The first two, already alluded to, are density and speed as manifested through dynamic pressure. From the figures shown, clearly the major perturbation of the magnetopause occurred between about 20 UT and 22 UT. During this period, the density was above average, and the speed well above average. The third factor is the orientation of the IMF, and in particular the northward component. During the major perturbation, as Figure 1.5 shows, this was solidly negative (i.e. the field was 'southward'). Under these circumstances 'reconnection' of interplanetary and planetary field lines can occur, stripping flux and plasma from the dayside magnetosphere and transferring it to the nightside [Hughes, 1995]. This process is likely of great importance in energy transfers within the magnetosphere.

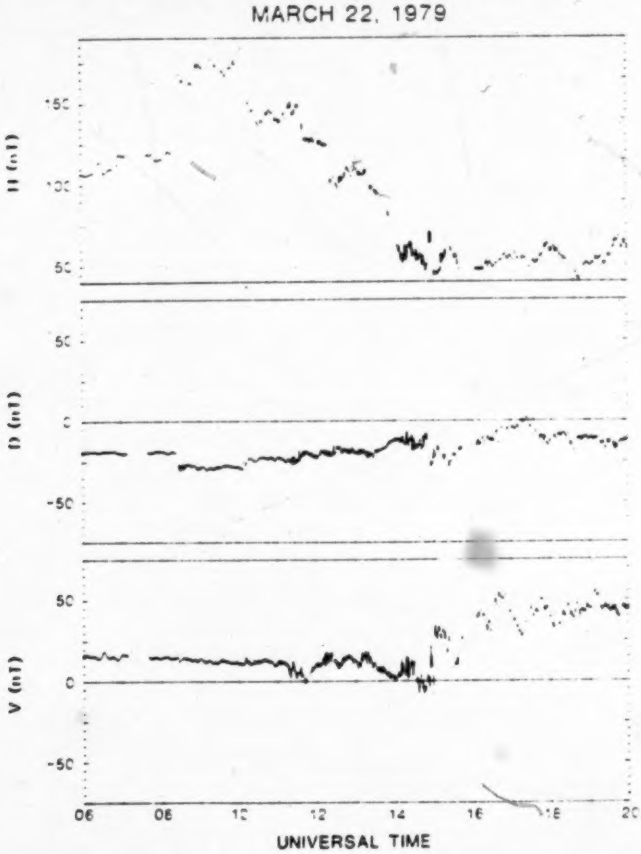
The example above uses satellite data to illustrate changes in the magnetopause. Ground observations also permit detection of buffeting of the magnetosphere by the solar wind. When its propagating shock waves bring pressure pulses into contact with the magnetosphere, compression results, with increase of the magnetic field inside. This may be seen from two points of view (as turns out to often be the case in describing magnetic fields in plasma). One is that the magnetic flux is essentially frozen in to the highly conducting plasma, so that simple hydrodynamic compression would lead to more flux in less space and thus a field increase. The other is that the magnetopause boundary currents (Chapman-Ferraro currents) strengthen to move the magnetopause boundary inward. In any case, a northward component increase is seen inside the magnetosphere when it is compressed. Such a compression is referred to as an SI (sudden impulse) or an SSC if it is followed by a geomagnetic storm. A very clear example is provided in the well-studied CDAW 6 event of March 29, 1979 [McPherron and Manka, 1985]. A shock in the interplanetary medium caused an abrupt increase in the solar wind velocity and density, and thus an increase in the dynamic pressure, at 0826 UT as seen near Earth. Simultaneously, ground stations worldwide recorded an increase in the northward component as seen in Figure 1.6.

Figure 1.6 Midlatitude H components on March 22 1979. SI at 0826 UT is evident while decrease at 1008 is less so. Substorm onsets at 1055 and 1436 are marked by vertical dashed lines and are accompanied by mid-latitude H bays.



The GEOS 2 synchronous spacecraft, in the vicinity of noon, detected an increase in the northward component from a relatively normal 120 nT to 170 nT in a matter of minutes at essentially the same time as shown in Figure 1.7. Coincident with the passing of a shock-associated current sheet at 1008 UT, the density abruptly decreased by a factor of about 3 to its pre-0826 value of about 10 particles/cm^{-3} (data not shown). The resultant decrease in dynamic pressure led to an immediate decrease in the northward component at GEOS 2. Signals on the ground also show this decrease, although it is not as clear as at the satellite. This behaviour clearly illustrates the effects of solar wind in controlling the size of the dayside magnetosphere. The 1008 UT current sheet passage also turned the interplanetary magnetic field southward, and this led to substorm activity in the magnetosphere, of a type which will later be discussed and seen to be characteristic.

Figure 1.7 · Magnetic field observations at the GEOS 2 synchronous spacecraft on March 22, 1979. The H (northward) component responds to compression of the magnetosphere by increased solar wind dynamic pressure at 0826 UT, while the other components are affected less. Pressure decrease at 1008 allows the northward component to decrease. Figure from McPherron and Manka [1985].



The effects noted above are associated with relatively rapid changes in the solar wind Recurrence of magnetic activity also follows the synodic rotation period of (near-equatorial regions of) the Sun, which is 27 days. This indicates that the causes of magnetic activity in the solar wind are generated by long-lived regions, likely near the solar surface. Due to the influence of the north-south component of the interplanetary magnetic field, relative to the orientation of the Earth's dipole, there exists a seasonal effect also [Russell and McPherron, 1973]. On yet a longer timescale, it is well known that overall activity of the Sun varies periodically in an 'eleven-year cycle', which in fact is a 22 year cycle if the polarity of the Sun's reversing main field is taken into account. The number of sunspots varies dramatically through the eleven year cycle and the level of magnetic activity observed on Earth is roughly correlated with the number of sunspots although not as variable. The modern study of auroral phenomena may be considered to date from the coordinated studies during high solar activity in 1957, known as the International Geophysical Year (IGY) [Akasofu, 1964].

b. Ionosphere

The transitional region between the neutral atmosphere and space is known as the ionosphere. In this region matter is ionized to a significant but variable degree, the primary sources of ionization being photonic radiation and particle fluxes. Recombination processes in the ionosphere, primarily in response to excitation through precipitation of energetic particles, provide one of the more inspiring forms of evidence of the solar-terrestrial interaction: the light of the auroras. In the lower ionosphere interaction with neutral molecules is a significant factor and the ions are dragged along through the Earth's magnetic field. This creates a unipolar inductor and creates a corotational electric field around the Earth.

As a plasma in a magnetic field, the ionosphere may be characterized by a conductivity tensor whose characteristics are dependent on frequency. The properties of this tensor at low frequencies (low compared to particle gyrofrequencies or the ionospheric plasma frequency) are discussed in the following chapter. At higher frequencies, in the radio range, the ionosphere can be dispersive, birefringent, and absorptive, and these properties are of practical importance and interesting in introductory discussions (see Jackson [1975], chapter 7, for example). Here we will simply note that the ionosphere and its disturbances can be of practical interest. Further, the ionosphere may be monitored and studied at radio frequencies. At these relatively high frequencies, phenomena occur near the electron plasma frequency which result in reflection of incident waves. Since the electron plasma frequency is $\omega_e = \sqrt{\frac{ne}{\varepsilon_0 m}}$, where n is the electron density and the other quantities physical constants (e the electron charge and m the electron mass), this reflection frequency may be used to make a profile of electron density at various altitudes. The important physical parameter of conductivity is largely dependent on electrons, whose density may approach 10^{12} m^{-3}. The altitude profile of the electron density may be ascertained by timing returns of pulses at varying frequencies giving reflections, and is called an ionogram [Davies, 1966]. The same technique also works if a satellite directs pulses downwards and is then usually referred to as topside sounding. Detection of increases in electron density may also be done by passive means, that is by monitoring natural radiations from distant space and noting that they decrease when precipitation of energetic particles causes the electron density to increase. Such detectors are called riometers. Sounding rockets and satellites also provide *in situ* measurements of ionospheric composition and density.

Ionospheric structure, as revealed by the above and other techniques, is complex. It is useful to speak of layers, these being, from bottom to top, the D region found below 90 km, the E region between 90 and 130 km, and the F region above 130 km. As will be discussed in the next chapter, for practical purposes, a height of 110 km is used as the 'nominal' height of ionospheric current flow in discussing low frequency magnetic perturbations. This height is used in modelling, with further justification below.

A further role of the ionosphere in magnetospheric studies is as a source or sink of particles. Since the Earth and Sun have differing cosmic histories, their bulk composition is different. In particular, Earth's atmosphere is abundant in oxygen and nitrogen, while the Sun, and thus the solar wind, is more abundant in hydrogen and helium (present in isotopic ratios not typical of Earth). Composition analysis can thus be useful in determining the sources of magnetospheric plasma.

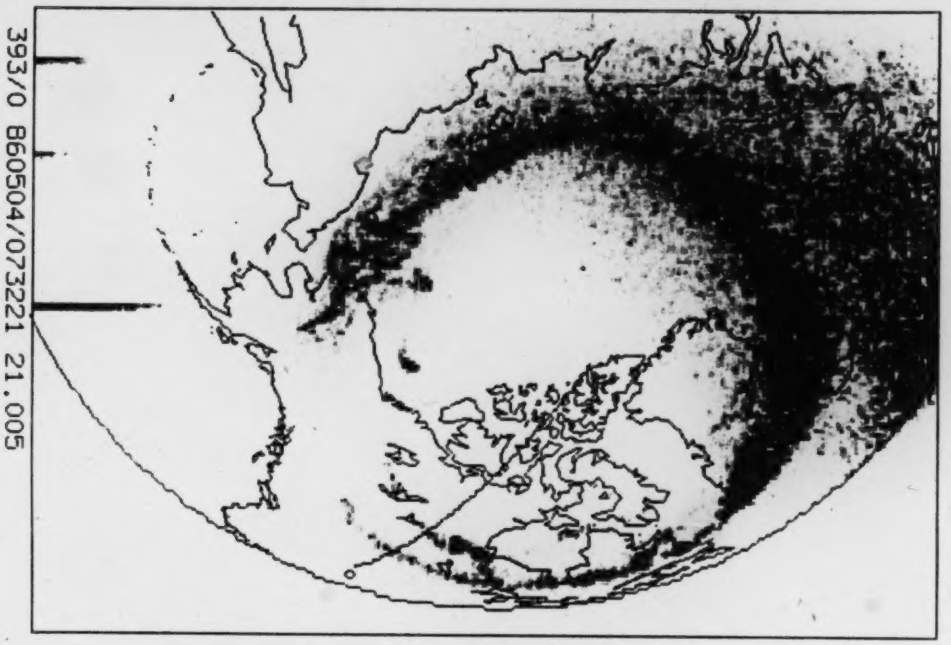
c. Magnetosphere-Ionosphere Coupling

Evidence of the interaction of the magnetosphere with the ionosphere is provided by the auroral emissions which result from the interaction of particles, accelerated in the magnetosphere, with the upper atmosphere. In addition to energy coupling by this mechanism of particle precipitation, there is electromagnetic coupling since currents flow through connected regions of the ionosphere and magnetosphere. Thus the two main types of Earth-based observations of the magnetosphere-ionosphere coupling have relied on detecting photons arising from the particle precipitation, or on detecting the magnetic fields associated with the currents. The currents are of course in most cases associated with particle precipitation in some way.

Both auroral emissions and magnetic perturbations are observed to be most pronounced in belts roughly centred on the magnetic poles of the Earth (see Figure 1.8), but usually about 20° equatorward. The large region where these phenomena are observed varies according to the level of the solar-terrestrial interaction, with the region where they are commonly observed being referred to as the auroral zone. The part of this region which is at any one moment active is referred to as the auroral oval as this is statistically, and for the most part instantaneously also, its shape. The most active part of the auroral oval is found near local midnight and the midnight sector is offset by at least 3° anti-sunward with respect to the magnetic pole [Akasofu, 1968].

Stations at typical auroral zone latitudes of about 65° in the magnetic dipole system are usually in the oval and detect activity most strongly around local midnight. Stations at higher latitude may cross the oval twice per day, as it remains relatively fixed with respect to the Sun-Earth line while Earth rotation carries them under it. In part due to a dearth of observatories, the highest latitude regions poleward of the auroral zone have historically been less observed than the zone itself. Even in the auroral zone investigations are limited by an insufficient number of observatories.

Figure 1.8 Instantaneous view of the entire northern auroral oval from the Viking spacecraft. Continental outlines have been superposed. Magnetic midnight is marked by a curved line through western Canada. Bars at left are imaging artifacts.



The magnetic effects of auroral currents have been used in the night sector of the auroral zone to trace the average position of the auroral oval [Rostoker and Phan, 1986]. These effects arise from electric current flow in the ionosphere in what is referred to as an electrojet, and the electrojet positions are within the auroral oval. The auroral oval is generally fixed in a sun-oriented coordinate system, that is, with respect to magnetic local time (MLT). Using data from a set of observatories along a meridian of geomagnetic longitude, the overall configuration of the auroral oval may be determined, since rotation of the Earth carries the observatories under it to sample all magnetic local times. As shown in Figures 1.9 and 1.10, the auroral oval borders traced in this way vary both with local time and with activity level.

Figure 1.9 Borders of auroral oval current flow in one-hour averaged magnetic data [Rostoker and Phan, 1986]. Dashed lines indicate borders of westward current flow, solid lines indicate borders of eastward current flow. AE 50 designates a period of globally low auroral oval current. Latitudes are referred to the magnetic pole.

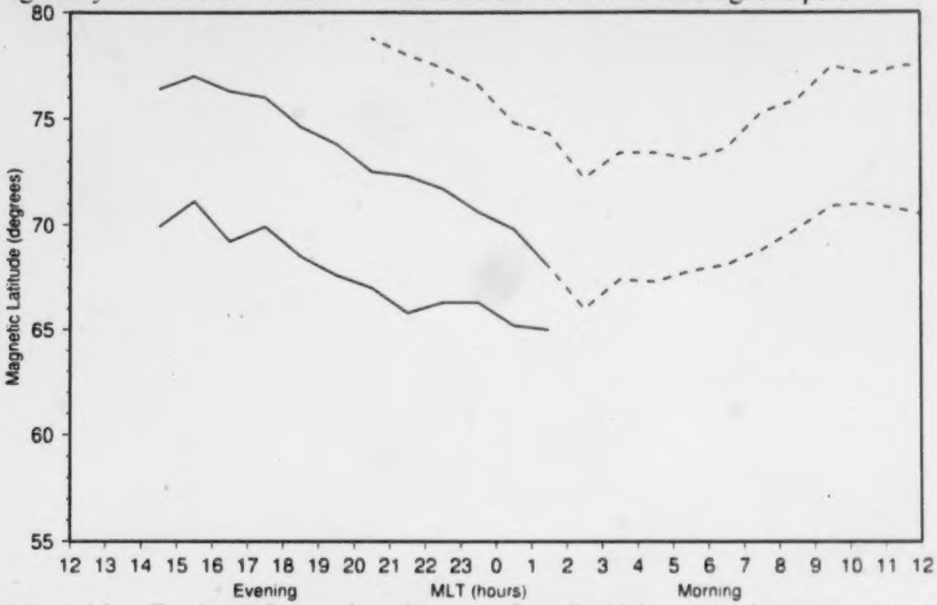
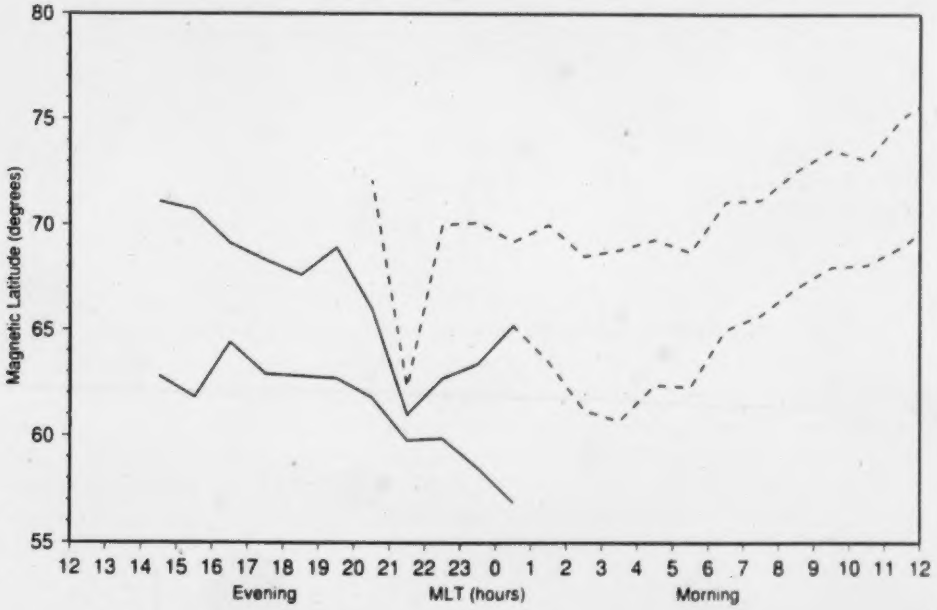


Figure 1.10 Borders of auroral oval current flow for high activity level [Rostoker and Phan, 1986]. Currents in this case are roughly ten times those in the previous figure. Evening sector electrojet in particular has shifted equatorward.



The evening sector is typified by an eastward flow of overhead electrical current at the lowest latitudes, with westward flow poleward of it as the midnight sector is approached. The morning sector has primarily westward current flow. Both in the morning and evening sectors, there is a decrease in latitude as midnight is approached. The overall pattern shifts equatorward by up to nearly 10° with increased activity, with most of the shift being in the evening sector, as seen in figure 1.10. Generally similar results came from studies of ionospheric disturbance at substorm maximum [Besprozvannaya *et al.*, 1991] as a function of local time. This study included not only magnetic fields but absorption of radio waves as a means of detecting auroral activity.

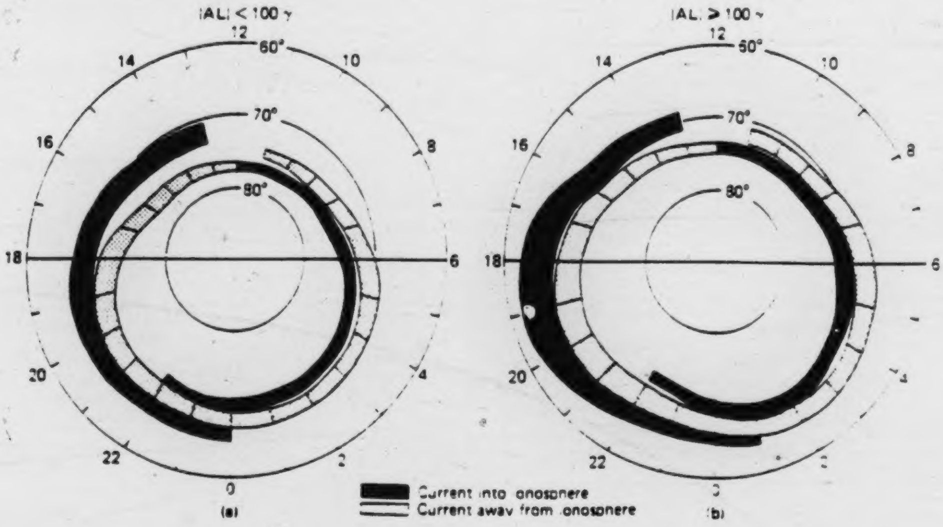
Since it is the aim of this thesis to achieve characterization of the auroral electric current systems with a simple model having few parameters, we should evaluate what we have seen so far. Organized current patterns are present in the ionosphere with a general pattern as described in the preceding paragraph. Our simple model should basically reflect these patterns provided current does not flow elsewhere. Since current flow is restricted to conducting regions and driven by electric fields, we must examine those aspects. The conductivity is discussed in more detail in the next chapter, but since it arises from photons or energetic particles impinging on the ionosphere, it is clear that it must be highest on the dayside and in the auroral zone where (recall Figure 1.8) particle precipitation is in evidence. The polar cap is the region inside the auroral oval and UV luminosity as in that figure suggests that there is little particle precipitation there. It could be suspected that photon-induced conductivity would be important in the polar cap when it is sunlit in summer, but various methods of measuring conductivity [Knipp, 1989, Chapter 4] indicate that the conductivity is low in winter and thus that little current flows there then. We will return to this point.

d. Configuration of the Magnetosphere

In this section the overall configuration of the magnetosphere is discussed. The actual causes of the near-Earth currents are the subject of ongoing research. Even the establishment of connections between near-Earth and more distant regions (primarily inferred through attempts to 'map' along field lines, but also deduced from particle distributions) is not generally agreed upon. Here an attempt is made to discuss accepted aspects of magnetospheric structure and the basis for this acceptance.

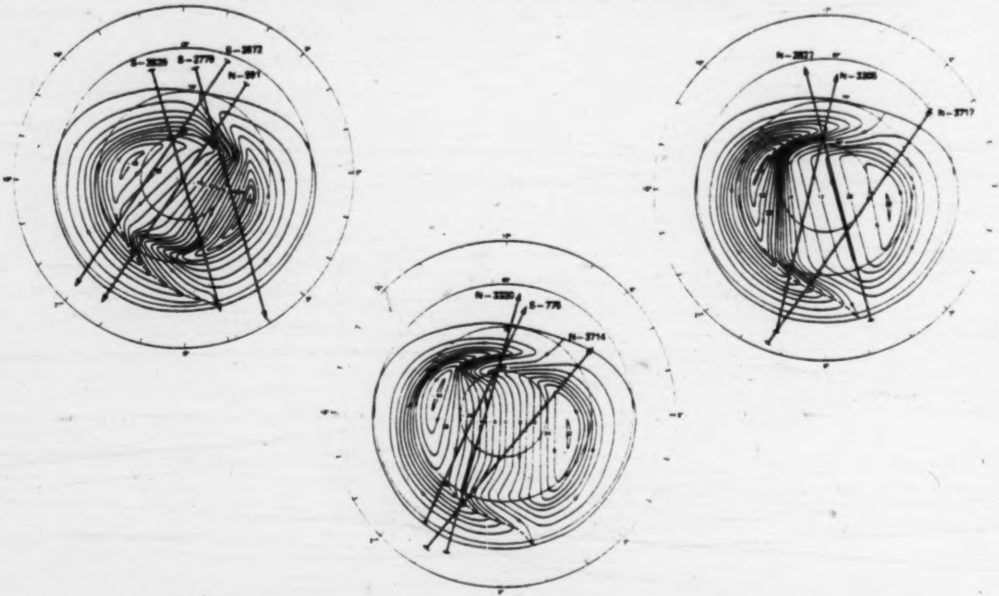
A major historical controversy in space physics concerned the structure of near-Earth space and provenance of auroral particles. As will be discussed in later chapters, magnetic observations from the ground do not allow unambiguous inversion to a three-dimensional causative current system. Despite the seemingly obvious connection between particle precipitation and near-vertical currents, space observations were realized to be needed [Boström, 1964] to distinguish between models of auroral electrodynamics with or without such currents. The initial relevant observations were done in the mid 1960s (see discussion in an historical review by Dessler [1983]), the currents were discovered, and within a decade the results were being systematized.

Figure 1.11 Canonical view of patterns of field-aligned currents observed with the Triad satellite [Iijima and Potemra, 1978]. Data from 439 passes during weakly disturbed conditions (left) and from 366 passes during active periods were averaged to produce the patterns shown.



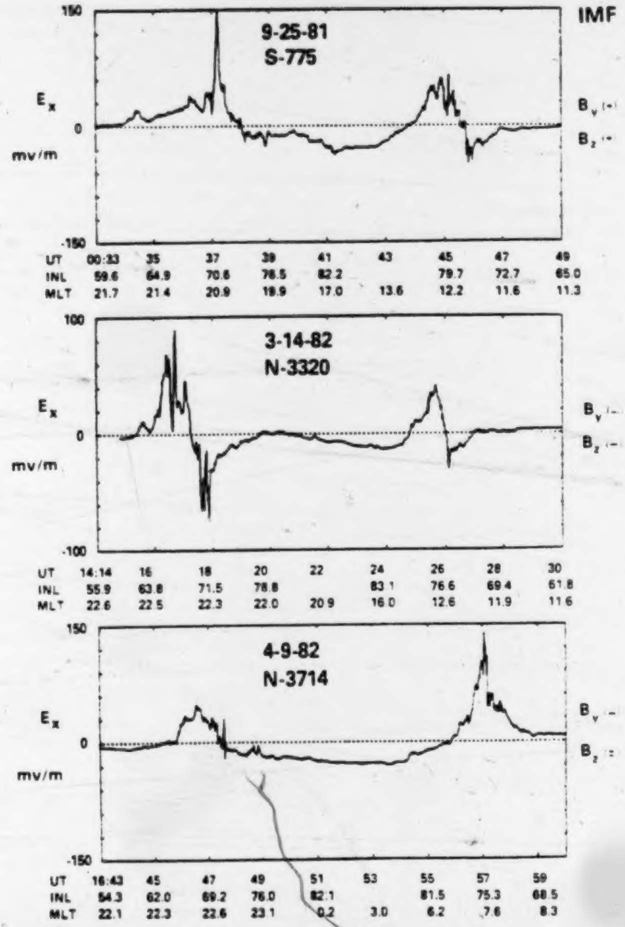
The establishment, through polar-orbiting satellite observations, that large-scale electric currents flow between the magnetosphere and ionosphere, allowed studies of those currents and their relation to other phenomena to be undertaken. Figure 1.11 shows the configuration of field-aligned currents which exist in a statistically repeatable fashion around the north polar regions of the Earth [Iijima and Potemra, 1978]. These currents were deduced from perturbations observed by the Triad satellite, in polar orbit at 800 km altitude. Interpretation of such perturbations is discussed in further detail in section 3.a.1.c, however here it will suffice to say that a satellite-borne magnetometer at 800 km basically allows determination of currents flowing at or near that altitude. Observations suggest that such currents largely flow parallel to the magnetic field, that is to say, are field-aligned. The basic configuration in both quiet and disturbed times is downward current poleward of upward current on the morning side, while the reverse is true on the evening side. In the noon and midnight sectors a more complex configuration is found. The poleward current zone is usually referred to as Region 1, while the equatorward currents flow in Region 2. Perturbations poleward of the well-defined Region 1 zone are sometimes referred to as Region 0 currents and do not have the same degree of statistical repeatability as those categorized by Iijima and Potemra.

Figure 1.12 Potential patterns as projected onto the northern hemisphere. Equipotential curves are marked with potential in kV. Model patterns vary with activity level and direction (sign) of azimuthal interplanetary magnetic field (Y IMF).



Observations of electric fields from satellites in low polar passes complement the magnetic studies. Examples of potentials and electric fields observed with the Dynamics Explorer satellite are shown in figures 1.12 and 1.13 [Heppner and Maynard, 1987]. Large fields and large field gradients are seen to be approximately colocated with the Region 1 and Region 2 currents in the auroral oval at roughly 60° to 70° invariant (dipole magnetic) latitude. From many passes the statistical picture of Figure 1.12 has been built up, in which the electric potential is presented. The electric field is the (negative) gradient of the potential, so that regions with closely spaced contour lines are those of high electric field. The units on the contour lines are kilovolts (kV) and it is seen that in this case there is an observed electric potential difference of roughly 80 kV across the polar cap. The models presented are for varying orientation of the toroidal (Y) component of interplanetary magnetic field. Here the similarities in patterns, rather than their differences, will be considered. The overall electric field within the polar cap is roughly from dawn to dusk, and 80 kV across the roughly 3000 km polar cap corresponds to an electric field of about 30 mV/m. However, within the auroral oval near dusk and dawn the potential gradient, and thus electric field, reverses. Near dawn the auroral oval electric field is equatorward while near dusk it is poleward. Individual spacecraft passes measured electric fields as shown in Figure 1.13. Reversals of the electric field correspond to reversals in plasma flow direction (see below) and this occurs in a pronounced fashion in the evening sector. This flow reversal had previously been found to occur based on studies of auroral arc drift and the dashed line shown to indicate it in Figure 1.12 depicts what is referred to as the Harang discontinuity.

Figure 1.13. Electric fields as measured on polar passes of the DE-2 satellite. Large fields and reversals occur in the auroral zone while the polar cap has small fields.



In a global sense, the above electric and magnetic observations are consistent in that ionospheric currents presumably, for the most part, connect the Region 1 and 2 currents to provide current closure. In the morning sector, downward currents are observed at the poleward border of the auroral oval, and the equatorward ionospheric electric field is that required to allow completion of the circuit to upward field-aligned current in the equatorward portion of the auroral oval. The situation is reversed in the evening sector. The more involved question of where the field-aligned currents close far from the Earth is not addressed here. It is simply worth noting that the auroral oval lies in the vicinity of the interface between closed field lines which can be traced from one magnetic pole to the other on Earth, and so-called 'open' field lines for which this is not possible [Lyons, 1992].

The electric fields are also consistent with what had previously been deduced about ionospheric plasma flow within the polar cap. The arrows on the equipotential lines in Figure 1.12 indicate plasma flow which would be inferred from the electric fields. The contour lines which indicate the equipotentials are also the plasma flow streamlines. This is due to the good conductivity of the upper F region ionosphere measured by the spacecraft at 800 km altitude, meaning that the 'frozen-in' condition applies: $\boldsymbol{E} = -\boldsymbol{v} \times \boldsymbol{B}$, where $\mathbf{E}$ is the electric field, $\mathbf{v}$ the plasma flow velocity, and $\mathbf{B}$ the magnetic induction. Since the electric field is also the (negative) gradient of the potential ϕ, $\boldsymbol{E} = -\nabla\phi$, its component along $\mathbf{v}$ is $-\nabla\phi \cdot \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{B} \cdot \boldsymbol{v} = \boldsymbol{v} \times \boldsymbol{v} \cdot \boldsymbol{B} = 0$. Thus $\nabla\phi \perp v$ and the equipotentials (orthogonal to the potential gradient) are locally parallel to the flow and are thus flow streamlines. The frozen-in condition applies throughout much of the magnetosphere (breaking down only on thin boundaries or small regions such as those above auroral arcs [Lyons, 1992]) and a further consequence is that the electric field is perpendicular to the magnetic induction (since $\boldsymbol{E} \cdot \boldsymbol{B} = -\boldsymbol{v} \times \boldsymbol{B} \cdot \boldsymbol{B} = -\boldsymbol{v} \cdot \boldsymbol{B} \times \boldsymbol{B} = 0$). This implies that magnetic field lines (the loci of connected points whose tangent lines are parallel to the magnetic induction at any point) are also equipotentials. A potential difference observed in the ionosphere will map out along the field lines and correspond to an electric field in the magnetosphere. This in turn indicates that flow must occur there. The corresponding magnetospheric and ionospheric flows are known as 'convection' by analogy to thermally driven flows in fluids (although the magnetospheric flows are not thermally driven).

The ionospheric flow associated with convection, provided that collisions with neutrals do not dominate ion motion (see Chapter 2), features ions and electrons moving at the same velocity [Carlson and Egeland, 1995]. In this case, the plasma velocity and the species velocities (indicated by subscript e for electrons, i for ions) are simply the '$\mathrm{E}\times\mathrm{B}$ drift velocity':

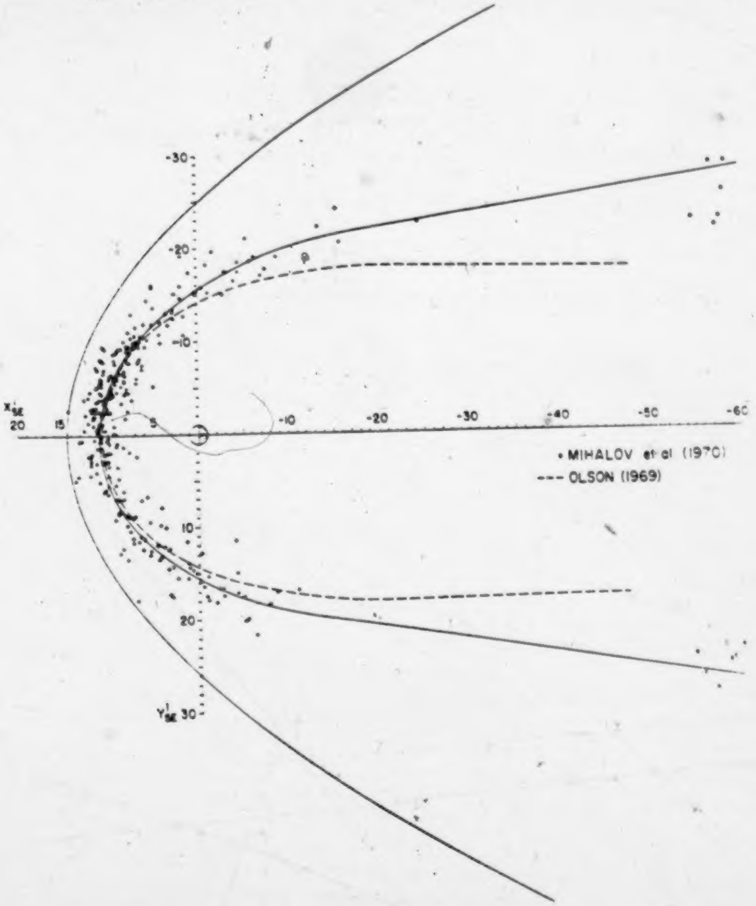
$$\boldsymbol{v} = \boldsymbol{v}_e = \boldsymbol{v}_i = \frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2} \approx \frac{|\boldsymbol{E}|}{|\boldsymbol{B}|} \boldsymbol{e}_\perp,$$

where the approximation at the right applies to horizontal flow in a near-vertical magnetic field and is parallel to the unit vector perpendicular to both. In the northern hemisphere that unit vector is parallel to a 90° counter-clockwise rotation of the $\mathbf{E}$ field. This approximation, which holds well in the auroral zone, is often used to allow conversion of drift velocities of ionospheric irregularities, which are readily observed with radars, into electric fields. With the known values of magnetic and electric fields in the auroral zone, the speeds of plasma flow along the ionospheric streamlines can be calculated. The behaviour of the main near-dipolar field at ionospheric height is given by equations below, but here it suffices to note that the auroral zone field is nearly vertical with a strength of about 50,000 nT. With the average 30 mV/m cross polar cap electric field, the associated plasma speed is 600 m/s. In regions where the electric field is stronger than this average the speed can be higher. About 1 km/s could be considered a nominal auroral zone ionospheric flow speed under active conditions. These considerations apply well in the upper F region of the ionosphere, where Carlson and Egeland point out that recombination lifetimes for ions are also long enough that these flows cause net transport of ions. In the E region, due to frequent ion-neutral collisions, the drift approximation breaks down, ions are not transported in this manner, and resistive effects are important. Discussion more appropriate to that region (which is, as mentioned above, where most current flows) is deferred to Chapter 2.

As mentioned in Section 1.c, the question of whether the polar cap carries current is important to this study. Certainly an electric field is present in that region as has just been shown. However, Iijima and Potemra's results, as discussed above, did not lead to a conclusion that there were systematic polar cap field-aligned currents present, except in the noon sector where they were extremely variable [Friis-Christensen, 1984]. Such a lack of field-aligned current does not rule out the possibility of electric current flow at ionospheric level. This could be expected in the sunlit hemisphere, and Knipp [1989] presents data showing that the polar cap does not have high conductivity when dark, a not unexpected result. Friis-Christensen found that nightside polar cap perturbations showed effects from large-scale field-aligned currents, while the dayside appeared to be dominated by ionospheric Hall current flow. In addition, the polar cap is responsive to the IMF in a more direct way than are the electrojet current systems. Attempts here to simplify the basic picture of high-latitude current systems dictate that the possible polar cap system not be included, especially since most of the studies are done with most of the polar cap dark. The degree to which this is justified, and to which polar cap perturbations arise from the systems included, will be discussed later.

The configuration of the magnetosphere implies connections of one sort or another between its various regions. We have seen that increased solar wind pressure can push back the magnetopause so that the magnetosheath may be detected at synchronous orbit. Nevertheless, it appears that the basic structure is preserved. The existence of such structure, in this case in the form of a persistent boundary, is also deduced from spacecraft moving through the more normal position of the boundary at quieter times as shown in Figure 1.14. The position of the bow shock is also shown in the figure although the positions of the individual bow shock crossings are not shown. The structures seen in these repeated spacecraft passes are entirely consistent with those seen when the boundary was pushed back to synchronous orbit by increased dynamic pressure. When that occurs, the compression of an overlying region has caused deformation of an underlying one, but the overall topology has remained the same. The changes in boundary positions can be well explained by considering fluid dynamics with magnetic effects added. This interaction between clearly defined regions may be considered to be one type of connection between them.

Figure 1.14 Magnetopause positions as projected onto the plane of Earth's orbit as observed by six IMP spacecraft in 1963 to 1968 [Fairfield, 1971]. Solid curves are best fits to magnetopause (inner curve) and bow shock (outer curve). Dashed curve is a theoretical magnetopause position.



Another type of connection between regions is that based upon mapping along field lines. The importance of such mappings may be looked at from two points of view. From a macroscopic point of view, field lines may at most points be considered to be equipotentials for reasons (mainly high parallel conductivity) mentioned above. This has implications for the mapping of electric fields, whose strength between equipotentials is related to their distance from each other and thus is closely linked to the magnetic field geometry. The integral effect of the mapped electric fields is expected to be a cross-tail potential of roughly 80 keV since field lines from opposite points across the polar cap presumably map to opposite points across the tail. As we have seen, the velocity is related to both electric fields and magnetic fields due to the frozen-field condition, and we will return to discussion of velocity fields at the end of this section. From the complementary microscopic point of view (the guiding centre approximation), relatively low energy particles follow field lines closely. Thus field lines indicate where particles can easily move from one region to another. Since field line positions are influenced by local currents (including those flowing along the field lines themselves), mapping is not as straightforward as it might originally seem, and in fact will be to a large extent avoided in the present work. It is nonetheless important to realize that points in the ionosphere may be connected to more distant regions of the magnetosphere as viewed both from a macroscopic and microscopic point of view. Details of the connections are found in mapping along field lines. The lowest order approximation to the field is that of a dipole. External to the Earth, the dipole field is curl-free and not the result of local currents. The actual mappings are made by its superposition with fields caused by external currents. Near the Earth, the dipole field dominates and mapping along its field lines is similar to mapping along the total field. Further out (starting at about synchronous orbit) the other fields can equal or exceed the dipole field and dipolar mapping should not be considered accurate. On the nightside of the Earth the currents generally cause true mappings from a given surface latitude to arrive at the equatorial plane further out than would a dipolar mapping. On the dayside, due to compression, the opposite is true. Here the considerations in tracing a dipolar magnetic field line are briefly discussed. In the modelling work presented later, currents are assumed to follow dipolar field lines as described in the next paragraph. Ground data are used in this work and the field near the surface is dipolar to a high degree of accuracy. That part of the surface fields which is due to currents flowing along field lines is well approximated by treating them as dipolar.

In quite general fashion, a magnetic field line is, at a given time, the locus of points whose tangent is parallel to the local magnetic field. To lowest order the Earth's internally generated field is that of a dipole: in Cartesian coordinates with the dipole moment M_z along the z axis this has components:

$$B_x = 3xzM_z r^{-5} \quad B_y = 3yzM_z r^{-5} \quad B_z = (3z^2 - r^2)M_z r^{-5} \quad B = \sqrt{3z^2 + r^2}\,M_z r^{-4}$$

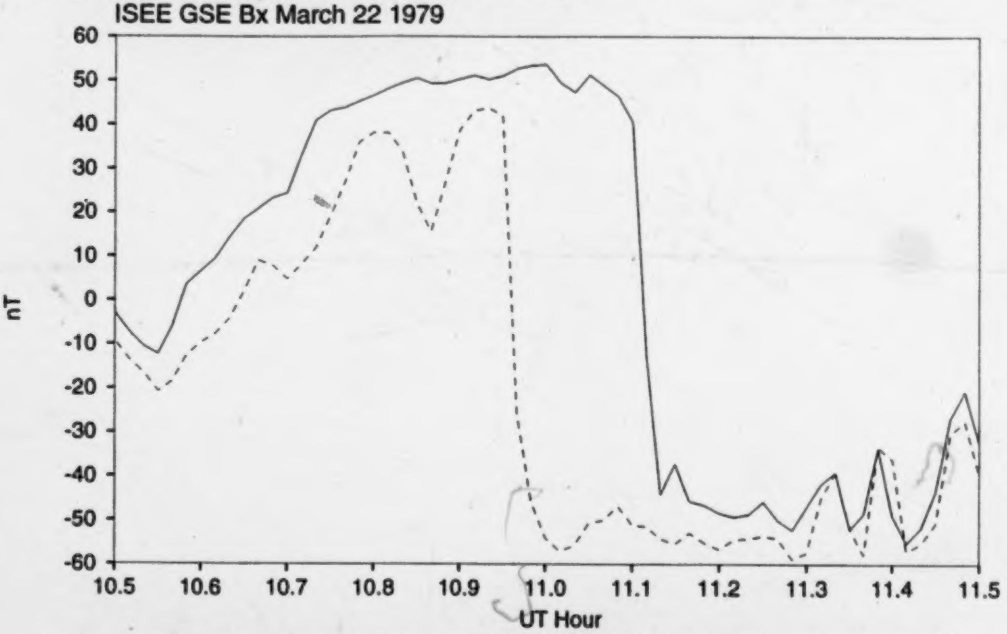
where $r = \sqrt{x^2 + y^2 + z^2}$. As noted above, for the terrestrial dipole the value of M_z is 8×10^{15} Tm^3 or 30400 nTR_e^3, the latter value being useful if distances are expressed in Earth radii (R_e) and magnetic induction in nanotesla (nT) as is usual in geophysics. Spherical coordinates are a useful system in which to express the equation of a field line [Parks, 1991, p. 55], which is $\phi = \phi_0; r = r_0 \sin^2\theta$. If the constant of integration r_0 is expressed in R_e it is often denoted by the symbol L and referred to as the MacIlwain L parameter of the field line. This is the radial distance from the centre of the Earth to the field line in the equatorial plane. A field line may be traced from the surface of the Earth, where $r = R_e$, and often it is convenient to use the polar angle of the start point as a parameter: calling this θ_0, one has $r_0 = \frac{R_e}{\sin^2\theta_0}$, or $L = \frac{1}{\sin^2\theta_0}$. In this case the solution for radius as a function of polar angle is $r = \frac{R_e \sin^2\theta}{\sin^2\theta_0}$ or in units of Earth radius $r = \frac{\sin^2\theta}{\sin^2\theta_0}$. Rather than the polar angle (colatitude) θ, it is often more convenient to use latitude as a coordinate, and in degrees the latitude $\lambda = 90^\circ - \theta$ Mapping in a dipolar field implies that field lines starting at latitude 45° have L values of 2, those starting at 60° have L values of 4, and those from 75°, L values of about 15. Typical auroral oval field lines should map into the equatorial plane of the magnetosphere at roughly 6 to 10 R_e. Magnetic field lines are distorted by the intensified currents which flow in active times and mapping is a hotly debated topic under these conditions.

The dominance of the internal dipole-like field is no longer complete beyond roughly 10 R_e. On the dayside, as already discussed, the magnetopause is at about this radius, and outside it one is in the magnetosheath or solar wind. On the nightside, the dynamic pressure of the solar wind creates the magnetotail (often referred to simply as the 'tail'). The rapid falloff of the dipole field means that the structure of the tail is dominated by local currents, with the dipole field ultimately forming a boundary condition. Consistent with this boundary condition is the observation that the northern half of the tail is dominated by fields which point Earthward[3] and northward, and that the southern half is dominated by fields pointing away from the Earth and also northward. Beyond 10 R_e downtail, the northward component is small compared to the total field except in the zone of transition between Earthward and anti-Earthward field. The field is referred to as 'tail-like' where this is true. The zone of transition is necessarily (through Ampère's Law) occupied by a current (oriented dawn-to-dusk) across which the field reverses. Since the field is small in this region the zone of current flow is referred to as the 'neutral sheet'. Magnetic pressure considerations dictate that lower field regions have a higher contribution from particle pressure than adjacent regions: the low field region in which the neutral sheet is found is thus called the 'plasma sheet'. The adjacent, higher field regions are referred to as the 'lobes' Figure 1.15 illustrates the oppositely directed lobe fields through simultaneous observation at twin ISEE spacecraft. The spacecraft are located only 0.6 R_e apart in the direction perpendicular to the location of the inferred neutral sheet, although at respective radial distances from Earth of 15.4 R_e and 13.9 R_e. ISEE-2 goes from north to south relative to the current at 10.95 UT, while ISEE-1 does not cross it until 11.1 UT. During this interval, ISEE-1 sees the Earthward-directed field of the north lobe while ISEE-2 detects the oppositely directed field of the south lobe. The small spacecraft separation allows inference of 'plasma sheet thinning' during the CDAW-6A substorm, March 22, 1979. In this case [Fairfield, 1984] it was inferred that the plasma sheet had become only 300 km thick (strictly speaking, using only magnetic measurements, this can only be inferred as the thickness of the current sheet embedded within the plasma sheet). During the short time interval shown in the figure, motion of the spacecraft is relatively small. In most cases, including this one, the crossing is not that of the neutral (current) sheet by the spacecraft, but a traversal over the spacecraft location by a moving current sheet. The term 'neutral sheet crossing' is often used when reversals of the tail field, as seen in this figure, are observed, but should not usually be taken to imply spacecraft motion.

---

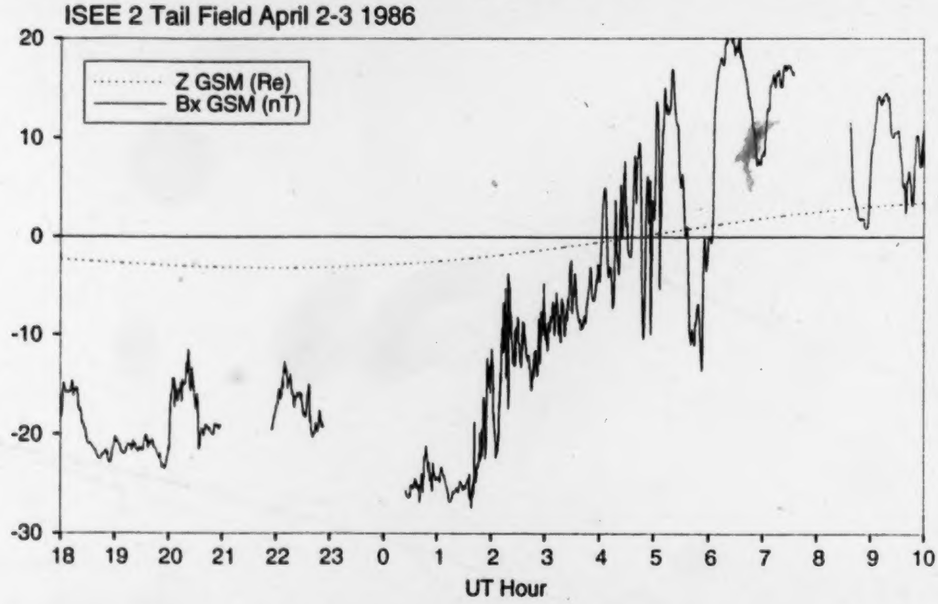
[3] Note that the main field as observed at the surface in the northern hemisphere is also 'Earthward', that is to say that its vertical component points radially inward.

Figure 1.15 Simultaneous observations of the two lobes from ISEE-1 (solid) and ISEE-2 (dashed). The GSE B_x component is along the Earth-Sun line and accounts for most of the field magnitude except when near zero.



A traversal through a less thinned current sheet is illustrated in Figure 1.16, in which data from the single spacecraft ISEE 2 are shown as its GSM Z position changes with time throughout an interval of sixteen hours. Initially, in the south lobe, the spacecraft detects a field directed away from Earth. This field is about -10 to -20 nT, smaller than in the CDAW 6 case above, partly because the spacecraft is further from Earth (at about 20 R_E) and the tail field is known to decrease with distance, and likely partly due to a lower strength of the current sheet. With much variation, the X component of the field reverses sign approximately as the spacecraft passes through the zero position of GSM Z, becoming about +20 nT on average after the crossing. At least part of the 'noise' in the plot is due to motion of the current sheet. This is best seen when the field briefly reverses for about one half hour starting at about 5 UT. The most likely explanation is that the mobile current sheet moved to positive Z GSM values. Thus, with respect to the current sheet, ISEE 2 was to the south and detected earthward field. This example also serves to illustrate the nature of spacecraft data taken at a single point. Even if structures (such as the current sheet) are present at all times, they may move and it can be difficult to disentangle their motion from that of a spacecraft. The reader may consult Section 3.e for details of the GSM coordinate system which provides a useful reference plane (Z=0) in this example.

Figure 1.16 ISEE 2 pass through the current sheet on April 2-3 1986. Z GSM coordinate very roughly indicates distance from the nominal current sheet position. B_x GSM is basically Earthward when negative. Data gaps are present in these data.



Since the energy density and pressure in a magnetic field are equivalent and equal to $\frac{B^2}{2\mu_0}$, energy is stored in the lobe magnetic field. Since the field is antisymmetric about the neutral sheet and zero at its centre, $B_x = \frac{\mu_0}{2} I_y$ in an infinite current sheet approximation, where I_y is the linear current density (A/m in SI units). An increased neutral sheet current thus leads to larger lobe fields and increased energy storage. This capacity for energy storage and the reversed field configuration, which is thought to be conducive to the reconnection process of energy conversion, have led to the concept of 'unloading' of tail energy in a bursty fashion. Much as discussed previously in the context of the upper ionosphere, the 'frozen-in' property of plasmas implies motion under the conditions in the tail. Thus energy may be stored in the form of plasma kinetic energy in the tail as well as directly in the magnetic field.

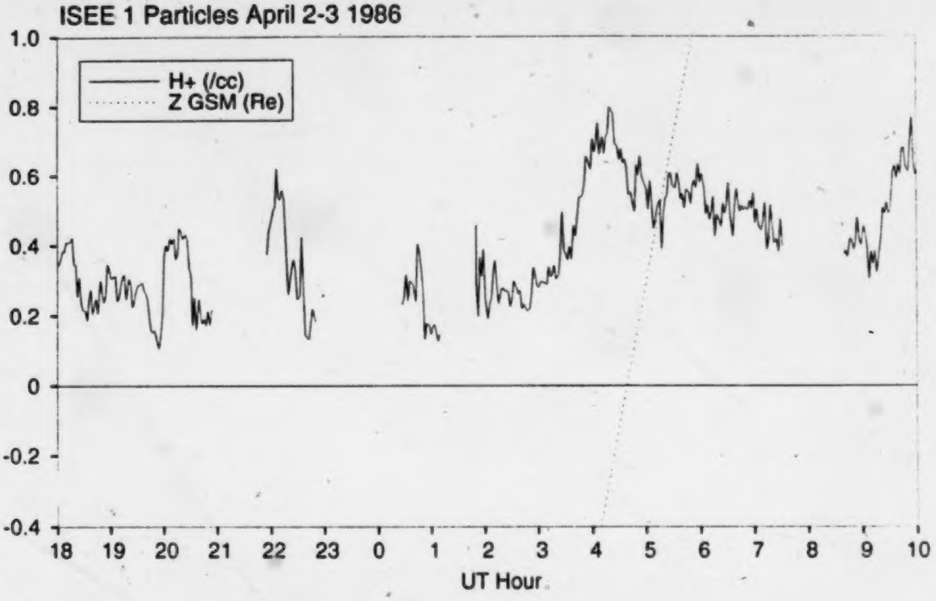
The final aspect of configuration to be dealt with here is the distribution of particle populations of the magnetosphere. As mentioned in the section on the ionosphere, it can be a source as well as a sink of particles. Characteristic species from that region (for example oxygen) can be injected into the magnetosphere to serve as tracers, and affect its structure and dynamics [Hughes, 1995]. Here, not much attention will be paid to the varying composition in the magnetosphere. The primary distinctions made will be of temperature and density of particles in various regions. It has already been established (recall Figure 1.3) that the total pressure in the magnetotail is much greater than the gas plus magnetic pressure in the solar wind, and can only be balanced by components of the solar wind dynamic pressure on its flanks. We have also seen that the north and south tail lobes, separated by a current sheet and associated plasma sheet, have oppositely directed fields. On the dayside, when the magnetopause was pushed earthward, another layer was noted, the magnetosheath. In fact, the magnetosheath plasma flows around the entire magnetosphere and has complex interactions with it [Hughes, 1995]. Inside the magnetopause there are boundary layers containing mixtures of magnetosheath and magnetospheric plasmas. The low latitude boundary layer (LLBL) is the region between magnetopause and plasma sheet. With the further subdivision of the plasma sheet into a boundary layer (between it and the lobe - this could be quite narrow) and a central plasma sheet, Hughes [1995] supplies typical parameters of these regions for the near tail (about 30 R_e out), which have been augmented by data from Hundhausen [1995] and Wolf [1995]) to form Table 1.1.

Table 1.1 Typical Near-Tail Plasma and Field Parameters

Region	n (m^{-3})	T_i (eV)	T_e (eV)	B (nT)	β
CPS	3×10^5	4200	600	10	6
PSBL	10^5	1000	150	20	0.1
Lobe	10^4	300	50	20	0.003
Magnetosheath	8×10^6	150	25	15	2.5
Solar Wind	7×10^6	10	12	7	2
Plasmasphere	10^9	1	1	1000	0.001

The last column indicates the plasma beta (β), which is the ratio of thermal to magnetic pressure. The relatively dense but cold plasma of the magnetosheath (density values cited here, in the tail, are considerably less than in the subsolar region) and the solar wind, and the less dense but hot plasma of the central plasma sheet (CPS), have high β and are dominated by thermal pressure, while the lobe is dominated by magnetic pressure and contains very few particles. The plasma sheet boundary layer (PSBL) is the transitional region between CPS and lobe and has intermediate characteristics. A distinguishing feature is sustained field-aligned ion and electron flows both earthward and tailward, indicating that acceleration processes are active in or near this boundary layer [Hughes, 1995]. Once more examining real data with all of its nonideal aspects, we consider Figure 1.17 which shows ISEE 2 particle data for the same period as the magnetic data of the previous figure. As the magnetic data suggested, the plasma sheet does not appear to be particularly thin since there appears to be no trace of the very low lobe densities in these data, even when the spacecraft is at ±2 R_E. The density appears to be higher when the spacecraft is traversing the neutral sheet (dashed trace for Z position crosses the axis), and certainly the density values are within the range indicated in the table for the PSBL and CPS.

Figure 1.17 Particle data as ISEE 2 traverses the neutral sheet on April 2-3 1986 Proton density (in units of 10^6 m^{-3}) increases near nominal neutral sheet (Z GSM=0)



As might be expected, particle behaviour in regions dominated by the dipolar field is somewhat different. Rather than a tail-like structure enforced by the streaming by of the solar wind, a cylindrical symmetry is dominant in the inner regions, imposed by the symmetry of the dipole field about the magnetic pole. In the innermost regions, called the plasmasphere, plasma is dense (up to about 10^{10} ions/m^3) and corotates with the Earth. Corotation of the bulk plasma may be regarded as a manifestation of the 'frozen-in condition' in the region where there is very low β so that the strong intrinsic field dominates. More correctly, in this region there is a corotation electric field established through charges (to first order) present at ionospheric level due to the rotation of that conductor enforced by atmospheric drag. Studies of dispersive electromagnetic waves ("whistlers") and *in situ* measurements from satellites reveal a usually abrupt plasmaspheric outer boundary referred to as the plasmapause [Wolf, 1995]. Here densities decrease by one or more orders of magnitude to those (10^6/m^3 or less) typical of the tail (CPS). In addition it has been observed that this boundary varies in position, being at lower L values for high levels of magnetic activity. At the highest levels it is at an L value of about 3.5 while in quiet times it retreats to about L=6. The plasmapause thus generally maps to subauroral latitudes.

Particles in the magnetosphere may be characterized by single particle equations, which for 'high energies' may give results differing from those of the bulk plasma. The guiding centre approximation of Alfvén is very suitable for examining the motions of such particles and a large and clear literature exists, of which two useful examples are Lyons and Williams [1984] and Wolf [1995]. Here only the general results of the theory, as needed to understand qualitative and some quantitative aspects of particle motion, are presented. The guiding centre is a fictitious moving point associated with a particle. Particle motion may be viewed as consisting of motion of the guiding centre and motion of the particle around the guiding centre in a circular orbit. Inherent in use of the guiding centre approximation is that the particle has free motion in the presence of external force fields which change on a scale much larger than its orbit about the guiding centre (both spatially and temporally). This is observed to be the case almost everywhere in the magnetosphere as generally the orbits are much smaller than any other length scales and the orbital frequency is one of the highest frequencies in the system. Frequent collisions (implying brief application of strongly varying forces) among particles would prevent the guiding centre approximation from remaining valid, but magnetospheric plasmas, right down to the upper ionosphere, are essentially collisionless. In this case, the effects of a force field **F** on a particle are to impart a velocity parallel to the field lines and to distort the orbits from circular: this latter causes motion of the guiding centre relative to the magnetic field. In a general sense, then, the motion of the guiding centre will take place at a drift velocity perpendicular to the local magnetic field [Parks, 1991, Section 4] which is $\boldsymbol{W}_D = \frac{1}{q}\frac{\boldsymbol{F}\times\boldsymbol{B}}{B^2}$, where q is the particle charge. A special case of external field is an electric field, in which the force on a particle is **F** = q **E**, and the charge cancels out in the drift equation, leaving the 'E×B' drift already alluded to in discussion of the ionosphere in the preceding section. Corotation may also be understood by considering all of the components of the plasma to E×B drift due to the induction electric field imposed by the ionosphere's behaviour as a unipolar inductor. Corotation applies to low energy constituents (the bulk of the plasma in the plasmasphere in practice) and we must now explain what is meant by low energy as opposed to 'high energy', the term with which this paragraph began. It has been noted that the cross polar cap potential is roughly 80 kV and that a potential of this order could be expected to map out into the magnetosphere. Particles having an appreciable fraction of the energy associated with acceleration through this potential could be thought of as 'hot' or 'high energy' since they would not change their energy appreciably in moving a small distance in the magnetosphere, while those of much lower energy would be expected to be more affected by electric fields (the observed cross-polar cap field, the corotation field, and local fields). The plasmaspheric particles, with an average energy of 1 eV, are most definitely 'low energy' and effectively react to the electric field through 'E×B drift'. Higher energy particles, although perhaps mixed with the cold particles, have significantly different dynamics. One should recall that these populations are not coupled by collisions and in many circumstances can be considered to be noninteracting. Exceptions to this statement can occur and can be important in generating magnetic pulsations [Samson, 1991]. We proceed now to consider more terms of the single particle drift equation, which are needed to understand high energy particles which will later be used in studying substorm onsets .

A particle (mass m) whose motion is controlled by the electric and magnetic fields is subject to the Lorentz force $\mathbf{F}_L$ and obeys Newton's second law, so that

$$m\frac{d\boldsymbol{v}}{dt} = \boldsymbol{F}_L = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}).$$

The orbits referred to above in the guiding centre approximation are part of the solution of this equation in a uniform magnetic induction field. It is readily shown that these orbits occur at the cyclotron frequency $\Omega_c = \frac{qB}{m}$, which is independent of particle energy, and with the Larmor radius (or gyroradius) $\rho_c = \frac{v_\perp}{\Omega_c} = \frac{mv_\perp}{qB}$, which increases with particle (perpendicular) energy ($v_\perp$ is the particle velocity along the orbit, which is in the plane perpendicular to the magnetic field). For all but the highest energy particles, and possibly near certain boundaries, the gyroradius is much smaller than other lengths in the magnetosphere. The electron cyclotron frequency is generally higher than other frequencies encountered, and due to inverse mass dependence is 1837 times higher than the corresponding proton (often loosely called simply 'ion') cyclotron frequency. Another aspect of the solution is that if there is a component of the electric field along the magnetic field, there is acceleration parallel to the field which is proportional to it. Empirically, the field lines are in most places very close to equipotentials, so that such an electric field is small. As seen above, the perpendicular component of the electric field causes an E×B drift, which may be considered as a fundamental type of particle motion.

A second type of motion characterizes particles in a nonuniform, and in this case, dipolar, field. Consider a particle in the equatorial plane in such a field and without any parallel component of the electric field, but having a possibly non-zero velocity ($v_\parallel$) parallel to the field. The particle's overall velocity vector will, at such a time, lie at an angle α to the magnetic field, where $\alpha = \tan^{-1}\left(\frac{v_\perp}{v_\parallel}\right)$. This angle is the 'pitch angle'. If the parallel velocity component is in fact non-zero, the particle will move away from the equatorial plane. In so doing, it will enter a region of stronger field (recall that $B = \sqrt{3z^2 + r^2}\, M_z r^{-4}$, which has a minimum in the plane z=0). In this motion, where the change in B will be taken to be slow over one gyroperiod, an adiabatic invariant applies [Wolf, 1995]. With a perpendicular coordinate's momentum-displacement integral conserved, the quantity $\frac{1}{2} m v_\perp^2 \frac{2\pi m}{qB} = \frac{2\pi m}{q}\mu$ is conserved, which means that, for constant charge/mass ratio, the so-called first adiabatic invariant $\mu = \frac{mv^2}{2B}$ also is. This quantity is equal to the perpendicular kinetic energy per unit of magnetic induction, and is the magnetic moment of the particle. In the small displacement considered here, the increase in B while μ is conserved means that the perpendicular kinetic energy must increase. Since it is easily shown that the magnetic field can do no work on a particle, its total kinetic energy does not change, so that any increase in perpendicular kinetic energy must be matched by a decrease in parallel kinetic energy. As the particle moves, due to its initial velocity, into regions of stronger field, it loses parallel energy (and thus speed) until it has zero parallel speed. At this point it reverses the process. The overall periodic motion between points of zero parallel speed is known as 'mirroring'. Where a particle mirrors depends on its equatorial pitch angle. If this is too large, in the case of the Earth, entry into the atmosphere may occur. Processes exist which can cause pitch angle changes ('pitch-angle scattering') and a population of mirroring particles may see its members enter the atmosphere if such processes are effective.

It is often stated that 'particles follow field lines' when the guiding centre approximation is valid. This is true of low energy particles, which in the general case make many 'bounces' and perhaps perform other motions on a much longer timescale. It should be borne in mind that even in the relatively simple geometry of the dipole field, particle trajectories are non-integrable. However, the problem has been solved repeatedly by numerical methods for even complex fields, and the following of field lines, with consideration of drifts perpendicular to them which are discussed below, is an established property. It can be shown for a dipole field that the general (any energy) case is for particles to orbit traversing the surface formed by rotation of the dipole field lines about the axis [Chamberlain, 1964, section 1.6]. In fields close to dipolar, closely analogous 'drift shells' may be defined using adiabatic invariants. For higher energy particles, in addition to the simple bounce motion, drifts may be important. The major drifts are the gradient and curvature drifts. The former is important for particles whose perpendicular energy leads to a large enough gyroradius that regions of varying induction field are sampled and the orbit deformed slightly. The latter is important for particles having large parallel energy, so that the action of following a field line implies a net inward centripetal force which is not negligible. By considering the orbital deformation from a general force **F**, the aforementioned drift velocity $\boldsymbol{W}_D = \frac{1}{q}\frac{\boldsymbol{F}\times\boldsymbol{B}}{B^2}$ results, even if we do not consider in detail the origin of the force but merely note that by following a drift shell the particle must have been subjected to a centripetal force (whose magnitude can be calculated based on the acceleration). The general drift formula in the absence of external non-electromagnetic forces thus becomes [Wolf, 1995] a sum of E×B, gradient, and curvature drifts, respectively:

$$\boldsymbol{W}_D = \frac{\boldsymbol{E}\times\boldsymbol{B}}{B^2} + \frac{W_\perp \boldsymbol{B}\times\nabla B}{qB^3} + \frac{2W_\parallel \hat{\boldsymbol{r}}_c \times\boldsymbol{B}}{qR_cB^2}.$$

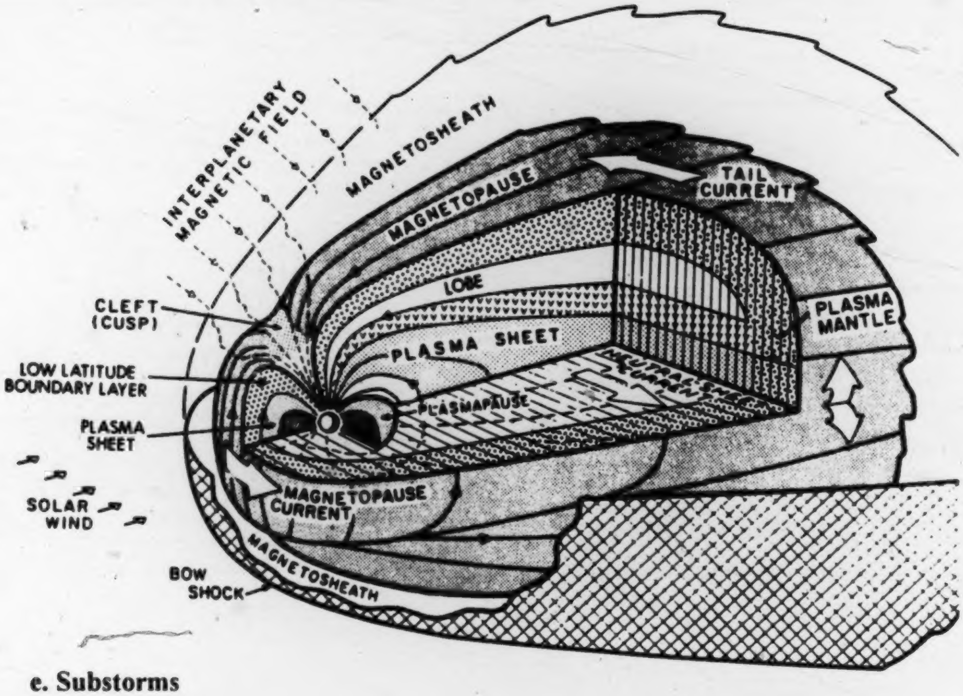
In this equation, $W_\perp$ and $W_\parallel$ are the perpendicular and parallel kinetic energies, respectively, and $\hat{r}_c$ the unit vector direction and R_c the magnitude of the radius of curvature. In a dipole field the field falls off outward while the radius of curvature is inward, so that the gradient and curvature drifts act in the same direction. This direction depends on the sign of the particles: electrons drift eastward while ions drift westward. For high energy particles, the second two terms are most important: their drifts are energy dispersive and capable of causing currents since opposite charges drift in opposite directions. During active times the ring current due to such drifts of high energy particles relatively near the Earth can contribute non-negligibly to the magnetic field observed at the surface.

As mentioned above, if the mirror points are low enough, bouncing particles may collide with particles in the atmosphere and not be able to return. This is known as particle precipitation and leads to excitation of atmospheric atoms and molecules, which upon recombination or deexcitation release radiation. The auroras observed at many wavelengths throughout the electromagnetic spectrum are one result of the precipitation. In a diagnostic sense, auroral emission may be used to locate areas of precipitation. Electric fields above the atmosphere may also remove atmospheric constituents (in a charged state), so that a process opposite to precipitation also occurs. Such upwelling would not usually be accompanied by emission but could carry net current. Study of upwelling ion beams with multiple spacecraft have established that under quiet conditions there is an auroral acceleration region which has a several kV parallel potential drop between about 1400 km and 2 R_e altitude [Reiff *et al.*, 1993]. In certain active regions the acceleration appeared to be at higher altitude, possibly in the magnetotail.

The low-altitude acceleration regions will not be discussed further here but lead naturally to the concept of acceleration in the magnetotail, which appears to have both continuous and impulsive components. The presence of magnetic and electric fields and plasma implies that there must be plasma drifts in the tail. Thus there is continuous motion. As discussed near the beginning of this section, this is also the case in the polar ionosphere and the flows are referred to as 'convection'. It stands to reason that there must be some connection between the ionospheric flows and those in the magnetotail: however this connection falls into the domain of mapping and here will not be considered to be known. Nevertheless, the structure of the tail and its flows can be consistently laid out.

This section has attempted to portray the structure or quasi-static configuration of near-Earth space with a strong observational basis. In closing, a pictorial synthesis of results is shown as Figure 1.18. The drifts of high energy particles alluded to above create a current which is not shown in the diagram: this is the ring current which is generally inside the plasmapause. The high energy particles are often referred to as the 'van Allen' radiation belts.

Figure 1.18 Cross-sectional pictorial synthesis of Earth's magnetosphere


INTERPLANETARY MAGNETIC FIELD
MAGNETOSHEATH
MAGNETOPAUSE
MAGNETOPAUSE
TAIL CURRENT
LOBE
PLASMA MANTLE
CLEFT (CUSP)
PLASMA SHEET
LOW LATITUDE BOUNDARY LAYER
NEUTRAL SHEET CURRENT
PLASMAPAUSE
PLASMA SHEET
SOLAR WIND
MAGNETOPAUSE CURRENT
BOW SHOCK
MAGNETOSHEATH

e. Substorms

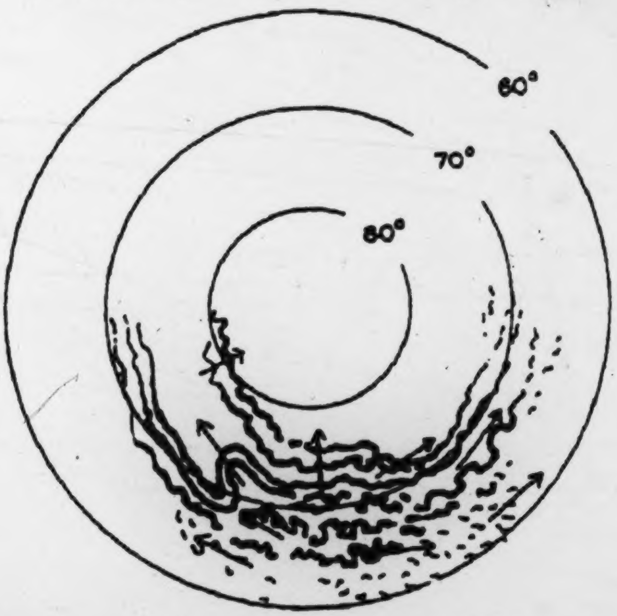
To this point, the focus has been on static characteristics of the coupled solar wind-magnetosphere system. Although it has been made abundantly clear that variations occur (for example in dynamic pressure), the overall picture which has emerged is one of regions of space with boundaries which may change but which nevertheless remain identifiable. By contrast, the substorm process appears to necessarily involve changes in the magnetospheric configuration, but it is not clear what these changes are, in which regions they occur, or to what degree the regions remain distinct.

Substorms are characterized by certain observational manifestations, which will be discussed here, some of which have relatively direct known connections to physical processes. The comprehensive physical picture of substorms, however, remains unclear, as the physical processes involved are complex and multifaceted, and the data available often ambiguous even when voluminous.

The term 'substorm' came into use around 1960, when Chapman and Akasofu, studying magnetic storms, found that the polar disturbance field varied on a shorter timescale [Stern, 1991]. In fact Birkeland in the period 1899 to 1903 had already identified what he called 'elementary polar magnetic storms', and made the association of increases in auroral luminosity with increases in magnetic disturbance locally in the auroral zone. The occurrence of substorms on the roughly half hour timescale and of storms on the half day and longer timescale is now accepted and the relationship between the two phenomena is still under investigation. Substorms can occur outside of storm periods but may be involved (in most cases) in producing particles causing storms. Storms are not discussed in detail here although their effect through the ring current will be seen below (Section 2.d) as manifested in the D_{st} index.

The various aspects of substorms led at one time to a set of terms describing their manifestations. It is now clear that all are related and thus the simple term substorm is in general use to describe the phenomenon. The systematic study of the visual auroral observational aspects of substorms on a near-global scale was first undertaken in the late 1950's [Akasofu, 1964]. Beginning from a quiet phase (it was later realized that there is usually characteristic activity during this phase also), in which homogenous arcs dominate the auroral zone, there is a sudden brightening of an arc and subsequent poleward motion. This 'expansive phase' leads to the formation of a region of intensified emission, usually centred on midnight, called the 'auroral bulge'. Figure 1.19 shows the later stages of formation of the bulge as inferred through ground-based imaging.

Figure 1.19 Stage III of the auroral substorm expansion phase according to Akasofu [1964]. Active auroras characterize the midnight sector (bottom).

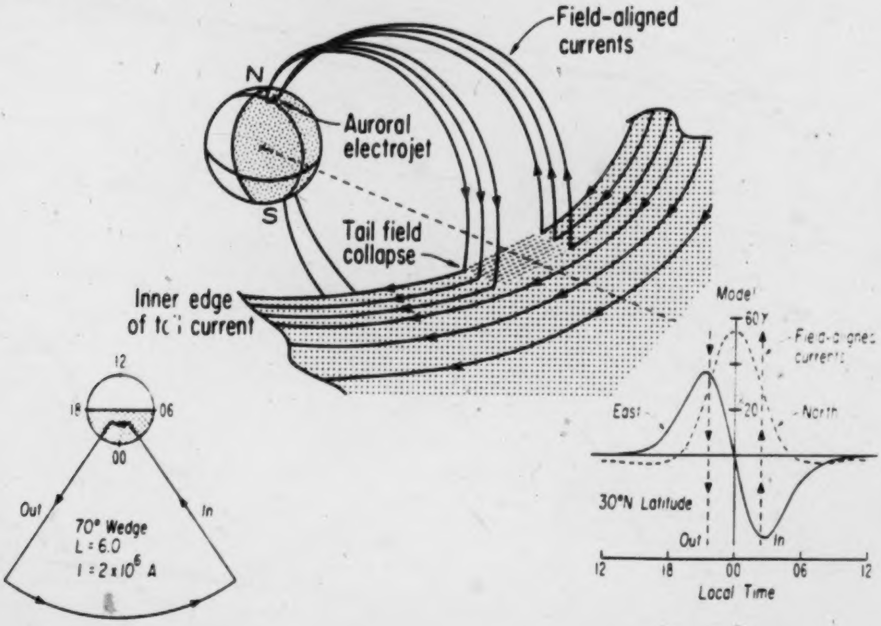


From instantaneous satellite imaging, it has been verified that the bulge is indeed roughly of the shape deduced by Akasofu (see particularly Chapter 7 below). A distinctive loop form in the evening sector is referred to as a 'surge' and subsequent more westerly appearance of surges on all-sky camera images at different stations led to the conclusion that the surge moved westward, hence the term 'westward travelling surge'. It has been subsequently shown that the westward motion is not necessarily smooth and that there may be more than one surge present at a time (these aspects are illustrated in Chapter 6 below). Seen in this figure but not emphasized by Akasofu is an evening sector band of enhanced luminosity at high latitude. A finding in the present study is that activity having substorm-like character can occur along such a high latitude band in the evening sector. Akasofu found that the maximum of the expansive phase was followed by the equatorward retreat of auroral arcs, with westward travelling surges degenerating into loops (including groups of loops according to Akasofu) and with eastward moving features known as omega bands characteristically found in the morning sector. This second phase of auroral activity was called the 'recovery phase'.

The expansive phase near midnight has as a very characteristic auroral zone signal a large (relative to other types of magnetic variation although small relative to the total field) decrease in the northward component of magnetic field. These signatures are often referred to as bays, and along with an analogous signature in midlatitudes, will be discussed below. Another signature of expansive phase onset is pulsational activity, particularly the Pi 2 damped wave trains in the 6.7 to 25 mHz band [Rostoker, 1979], observed in midlatitudes and increasing in strength near the auroral zone electrojet. Within the auroral zone, the northward magnetic perturbations of 'negative H bays' may show a very rapid change of up to 1000-2000 nT in a timespan of a few minutes. The Z (vertical) component can also show large changes but these are more complex as they depend critically on the latitude of the observer relative to ionospheric currents which are inferred to be enhanced in the ionosphere during a substorm. The east-west component near the auroral zone is also perturbed near surges in such a way as to suggest north-south current flow to be associated with them. Both ground and satellite data suggested that during a substorm there was formation in the near-midnight auroral zone of a region of enhanced ionospheric current, with field-aligned currents connecting it to somewhere in the magnetotail. Not only current closure, but also detection of the magnetic field of these field-aligned currents from the Earth's surface, supported this view [Clauer and McPherron, 1974]. That such a three-dimensional current configuration could exist was by that time tenable due to the discovery (in confirmation of Birkeland's hypotheses) of field-aligned currents, which had occurred by 1970 [Dessler, 1983]. This configuration, apparently active during the expansive phase, is called the 'substorm current wedge' and Figure 1.20 shows the canonical view of it.

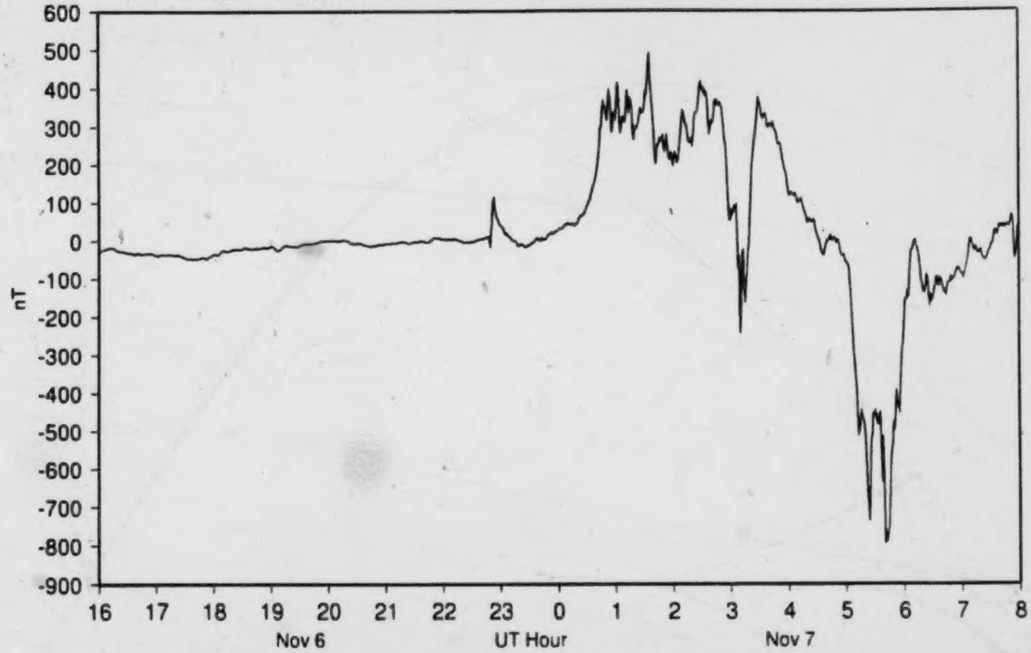
Some auroral zone geomagnetic bay variations are known to be associated primarily with the near-Earth substorm current wedge. Similar perturbations at midlatitudes may also be explained by the wedge. Studies in the 1960's indicated the primacy of the westward electrojet in the midnight sector during substorms and equivalent current systems of such an electrojet could explain them [Rostoker, 1966]. In the D (eastward) component, positive bay occurrence peaked strongly about 3 hours before midnight and that of negative bays about 2 hours after. Viewed with the modern knowledge of the importance of field-aligned currents to the mid-latitude D component, these statistics are entirely in conformity with the Clauer and McPherron view as presented in Figure 1.20. These midlatitude perturbations, appropriately weighted, can be used with a magnetic modelling routine to constrain the position of the ionospheric currents further poleward. Rostoker's statistical data also allows deduction of the rough size of the substorm current wedge as being 5 hours of local time or 75°, comparable to that given as typical by Clauer and McPherron [1974].

Figure 1.20 Canonical view of the substorm current wedge as deduced by Clauer and McPherron [1974]. Midlatitude perturbations suggest downward current east of midnight and upward current west of midnight.


Field-aligned currents
N
Auroral electrojet
S
Tail field collapse
Inner edge of tail current
12
18
06
00
Out
In
70° Wedge
L = 6.0
I = 2 x 10^6 A
Model
60γ
Field-aligned currents
East
20
North
30°N Latitude
Out
In
12
18
00
06
12
Local Time

The modern 'canonical view' of substorm phenomenology adds a third phase to the two noted by Akasofu. This is the 'growth phase' first clearly identified by McPherron [1970]. The initial motivation for the identification of this phase, preceding the expansive phase, was the presence of disturbances in ground magnetograms. A good example is seen in Figure 1.21, which shows the perturbations to the northward component during an intense event, as recorded at Meanook, 140 km north of Edmonton. An SI at 2300 UT on November 6, 1997 is followed by positive signatures in the early hours of the following day, while the station was in the evening sector.

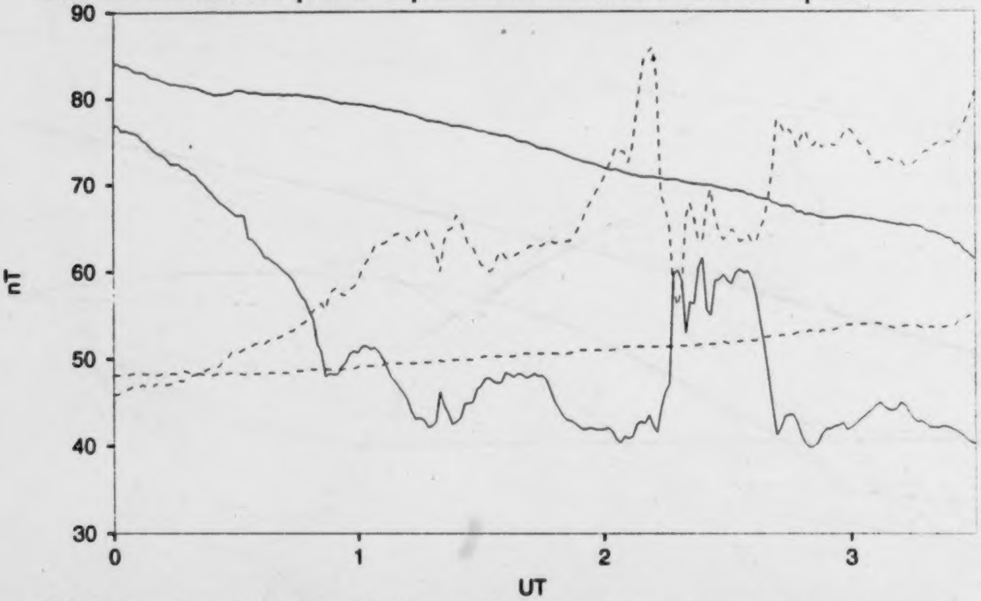
Figure 1.21 Northward magnetic perturbations at Meanook on November 6-7 (UT) 1997. An SI, growth phase, and expansive phase signatures are seen.



Expansive phase signatures of rapid decline in the northward component (bays) are seen at about 0300 and 0500. These primarily show the effects of the near-Earth electric currents of a near-midnight substorm current wedge similar to that in Figure 1.20. The positive perturbations cannot originate from this system and indicate a growth phase current system which was present in the evening sector. The early studies showed that in addition to ground signatures, distinct magnetospheric signatures were found to occur during the period before the expansive phase and these are now regarded as characteristic of growth phase (see below). It is interesting to note that bursts of Pi 2 band pulsations are taken as signals of the expansive phase, but McPherron [1970] noted that they are also seen during the growth phase. At that time it appeared that the growth phase signatures were generally of the same sign as those of the expansion phase and that a separate current system was involved. It is now generally thought that the substorm current wedge, subsequently proposed by McPherron as noted above, is in fact a separate current system, and that the growth phase signatures primarily arise from the 'driven' system [Akasofu, 1983]. The driven system is a pattern of currents on the global scale. Initially suspected in the 1930's, it has historically been referred to as the S_D system, DP or DS. In the view of Chapman the global effects seen were interpreted as ionospheric currents everywhere. The present three-dimensional view is that the effects in the auroral zone are primarily those of ionospheric Hall currents, while those at low latitudes are generally those of field-aligned currents. The basic configuration of the driven system is one of westward electrojets in the morning sector and eastward electrojets in the evening sector, consistent with what the perturbations in Figure 1.21 suggest. To create the global driven system, net field-aligned current input occurs near noon and return currents to the magnetosphere are near midnight These global scale systems may also have Region 1/2 currents associated with them, but the perturbations they cause on the ground are generally subtle. Here, it is shown that they can be detected and modelled, in the polar cap primarily. A discussion of the driven system in terms of relation to magnetospheric convection and electric fields is given by Lyons and Williams [1984, pp. 60-63], in which its relation to active conditions is emphasized (as this means it cannot be used to study general quiet-time convection). The driven system's net convection electrojets have also been referred to as the DP2 system (see for example [Kamide, 1988, pp. 671-672]) while the substorm electrojet is referred to as DP1. The term 'driven system' implies a physical mechanism and similarly the diversion of field-aligned current into the ionosphere in the current wedge model is referred to as an 'unloading' process. The relation of the driven and unloading systems is not yet clearly established and is a subject treated later in this work. A brief discussion of satellite observations of substorm phenomenology will now be given.

In the midnight sector at geosynchronous altitude (relatively near Earth in terms of position downtail) the magnetic field may be characterized as transitional between dipole-like and tail-like. The dipole-like internally-generated field suffers a 300 fold decrease in magnitude in going from the surface to 6.6 R_e in the equatorial plane, resulting in about 100 nT total field there, most of which is directed northward under quiet conditions. Tail fields are variable but can at times attain almost this value, but mostly in the radial component. Figure 1.22 shows the field detected at the GOES 5 spacecraft under quiet and disturbed conditions. The quiet day curves correspond very closely to those found for synchronous orbit at this time of year by Skone [1994]. The spacecraft is in the northern magnetic hemisphere, and thus the earthward component is positive, that is to say, the overall field direction is toward the Earth. On the disturbed day, the earthward component exceeds the quiet day value at all times after 0030 UT. From 0215 UT until 0245 UT there is a notable decrease in the earthward component, and a corresponding increase in the northward component, with both approaching, but not attaining, quiet day values. An increase in the earthward component corresponds to a more 'tail-like' field and departure from dipolar geometry. The decrease in earthward component and rise in northward component indicate a return to more normal dipolar conditions at synchronous orbit. This process of 'dipolarization' in the near-Earth tail occurs at the same time as many indicators of substorm onset, notably major ground perturbations in the late evening sector below the satellite (see Chapter 6). This example is consistent with the observations of Nagai [1982] using earlier GOES-series satellites in that substorm onset corresponds to dipolarization. Another well-studied example is the case of GOES 3, located at 0145 LT (i.e. near midnight) during the CDAW 6A substorm onset [McPherron and Manka, 1974]. That case is more dramatic than that shown here in that the earthward component decreased its magnitude from 100 nT to about 30 nT in only a few minutes at onset time. In the event presented here, the dipolarization is temporary, whereas the events studied by Nagai were selected to be simpler and to show a long-lasting dipolarization (as was the case for CDAW 6A also). The data here support the conclusion of Nagai: that there is a current flowing in the equatorial plane near synchronous orbit which initially produces a more tail-like, distorted configuration, and that that current is disrupted by substorm onset. The large perturbations observed on the ground during this disruption have led to the view that the current is diverted to the ionosphere as suggested by the canonical current wedge as pictured in Figure 1.20.

Figure 1.22 Northward and earthward magnetic components at GOES 5 on April 3 and 4 (smoother curves) 1986. Solid lines indicate the northward component, dashed lines the earthward component. Dipolarization is visible at 0215 UT April 3.



Substorms are also observed to accelerate particles within the magnetosphere. Phenomenologically, such acceleration is detected through flux increases observed at or near the beginning of the expansive phase. While some flux changes are due to changing magnetic fields and their influence on particle motions, overall enhancements of the flux levels above previous levels also occur. Observations of drifting populations and ion beams in regions near the plasma sheet boundary layer indicate that acceleration actually occurs and appears to be linked to induction from rapidly changing magnetic fields [Lopez and Baker, 1994]. Further, this occurs in the tail at roughly 10 R_e downtail, and up to several hundred keV of energy may be involved, although the bulk of the plasma remains at far lower energies. This acceleration is quite distinct, in location and energy, from that which occurs at much lower altitude above auroral arcs. Particles traversing the lower acceleration region are typically accelerated to 1-10 keV, and account for brightening of the auroras at expansion phase onset. Aspects of the use of both high and low energy particles in studying substorms will be discussed in Chapter 3. Here it will suffice to note that a large particle data set exists with which substorm injections may be studied [Baker *et al.*, 1982], and it has been amply demonstrated that sharp flux increases are associated with substorm onsets.

Figure 1.23 Electron fluxes in the 30 to 300 keV range at synchronous satellite 1984-037. Lowest energy is at top in each set of curves. UT is marked on bottom axis, local time on top horizontal axis.

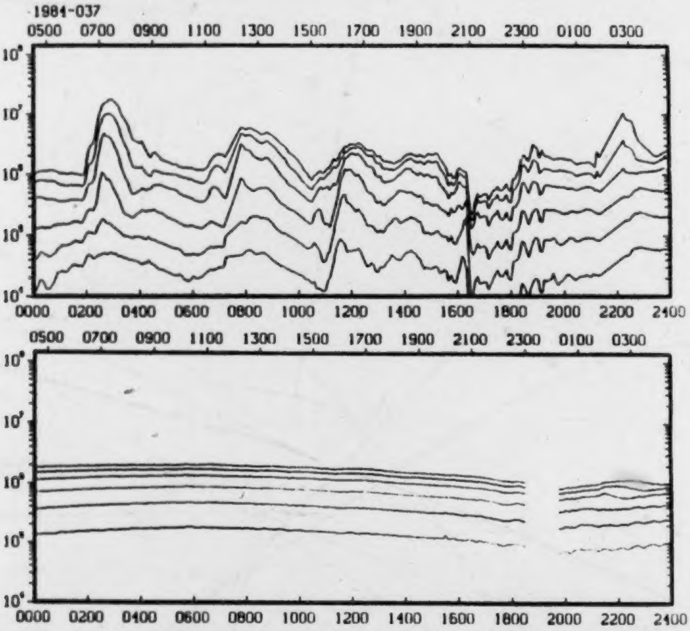


Figure 1.23 shows part of the large synchronous orbit dataset and compares fluxes on a quiet day with those from a more active period. In general, on the active day the highest energy fluxes are depressed and the lowest energy fluxes experience large transient increases. Such behaviour is typical of substorm injections, and with several spacecraft the energy dependent drifts allow determination of the injection region. In this example a flux dropout which is likely due to a change in magnetic field configuration is seen at roughly 1620 UT. The impulsive flux increases at earlier times would most likely be associated with injections. Further examples of injections associated with expansive phase onset may be found during the major substorms studied in later chapters.

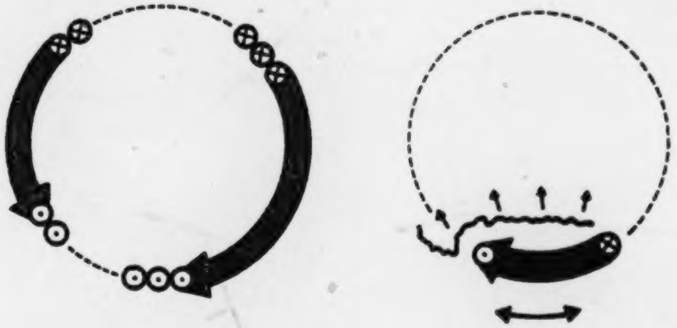
This section has briefly reviewed aspects of substorms, which appear as transient events within the larger magnetospheric context presented earlier in the chapter. We now examine the essential characteristics of electric current systems that flow during substorms, which form the basis of global modelling undertaken later in the thesis.

f. Electric Current Systems for Substorm Models

In Section 1.c above, basic aspects of magnetosphere-ionosphere coupling have been detailed. Figure 1.8 illustrated the auroral oval, and the results of Rostoker and Phan [1986] in outlining its average position through observations of ground magnetic perturbations were given. Ground observations form the basic data for modelling in this thesis. As was alluded to above, the ground perturbations reflect the positions of electrojets, that is, regions of electric current flow in the ionosphere. It will be shown in Section 4.d that two basic types of current system exist and that they are to a large degree colocated. Systems whose ionospheric currents are largely perpendicular to the trend of the auroral oval, such as those associated with the field-aligned currents whose patterns are shown in Figure 1.11, do not produce large perturbations at ground level. Systems with large longitudinal extent and current flow basically parallel to the auroral oval do produce relatively large perturbations on the ground. From observations of these perturbations, a general picture of three main current systems has been built up, and is shown in Figure 1.24. In this diagram the leftmost sketch shows schematically the three-dimensional current system associated with the electrojets usually present in the ionosphere, described above as DP2 or the 'driven' system. Downward net field-aligned current (shown by circled crosses) on the dayside feeds ionospheric currents flowing toward the nightside, where it leaves the ionosphere as upward field-aligned current (shown by circled dots). This results in large regions of nearly meridional electric current flow in the ionosphere which account for the overall pattern deduced by Rostoker and Phan. The figures showing their results also give an idea of the latitudes at which the large-scale currents flow. Generally there is eastward current flow over the evening sector of the Earth, and this gives rise to a northward perturbation at ground level, as was noted in association with Figure 1.21. The morning sector is typified by westward current flow and southward perturbations at the surface.

On the right hand side in the figure, the current associated with the substorm current wedge is shown, once more with a region of downward current shown as a circled cross and a region of upward current as a circled dot. The wedge current may be compared to the canonical view of Figure 1.20. The dynamic development is indicated by small arrows indicating poleward motion and east-west spreading of the region of current flow. These are generally similar to the arrows indicating the development of visual auroras as seen in Figure 1.19. The two types of systems shown are felt to schematically represent the patterns of electric current present in the auroral zone. The degree to which one or both types of current system are present, and the relation between them, are subjects of active research. The morning and evening system electrojets forming the 'driven system' appear to exist independent of substorm activity although they certainly are involved in substorms as cursory examination of magnetometer data has already shown (refer once more to Figure 1.21). As discussed, the current wedge is one of the most prominent features in substorm development.

Figure 1.24 Simplified ionospheric current systems for substorm models. Modified from Kamide and Baumjohann [1993].



As mentioned in the Preface, the mechanisms behind substorm development remain obscure. It is the intent of this thesis to characterize the near-Earth current systems in terms which allow this development to be better studied. By casting the model in terms of identifiable systems, analysis, a process of breaking down the problem, becomes possible. If well-characterized near-Earth systems reflect different processes in the magnetosphere, that characterization will clearly aid in understanding the physical bases of those processes. The ties inferred between components of the simple sketch illustrated above and the driven and loading-unloading processes make it attractive as a basis for modelling. Quantitative investigation of such models is the aim here.

To the extent that other current systems may be neglected (a point considered in assessing results obtained), it should be possible to use models based on three-dimensional representations of the driven system currents and the wedge to reproduce the magnetic data observed during active times. The present study takes this approach by using a modelling routine with three basic current systems, and attempting to vary their parameters until a good match to the input data is found. The current systems parametrized in this manner constitute an inversion of ground magnetic data in terms of realistic physical sources. The feasibility and utility of this approach to inverting magnetic field data associated with substorms will be evaluated in what follows. After further discussion of relevant physics and instrumentation (Chapters 2-3), as well as of other techniques allowing inversion of observations of magnetic field perturbations into representations of causative current systems (Chapter 4), a new technique allowing direct inversion of the Biot-Savart integral in three dimensions will be presented (Chapter 5). This technique, Automated Forward Modelling (AFM), will be used in five event studies (Chapters 6-10) for which good data are available. Comparison of AFM results with other information about the events allows deduction of numerous aspects of auroral phenomenology. AFM's advantages, including the fact that it may be used globally or regionally as appropriate, will be exploited, and its place among techniques for studying the Solar-Terrestrial interaction will be evaluated.

Chapter 2. Physical Processes

The previous chapter has considered the structure of near-Earth space and the interplay of that structure with the solar wind. In this chapter some of the physical processes operative in near-Earth space are discussed, mainly with the intention of placing in context the study of geomagnetic variations occurring with a timescale of a few minutes. Pulsational phenomena on shorter time scales are not considered in detail in this study and the physics behind them is thus also not discussed here.

a. Physical State of Near-Earth Matter

Above an altitude of roughly 90 km the tenuous gasses which fill near-Earth space have a significant degree of ionization which generally increases with height. The collision frequency is low, which does not favor recombination. On the other hand, there are abundant sources of energy for ionization, including solar photons and high-energy particles of magnetospheric, solar, and cosmic origin. At ionospheric height solar photons dominate in maintaining ionization on the dayside, but are not available to do so on the nightside. The auroral zones, where ionization is maintained by particle precipitation, have an ionospheric conductivity roughly equivalent to that on the dayside while other regions of the nightside have lower conductivity due to the lower ionization. The actual day/night ionization variation outside the auroral zone is roughly a factor of 100.

The most important aspect of the presence of ions is that the medium becomes conducting and subject to electromagnetic forces. The dynamics of such an ionized gas, or plasma, are more complex than those of a neutral gas. In lower regions of the ionosphere where there is still an appreciable fraction of neutral molecules, their interaction with ionized constituents may be important.

b. Magnetohydrodynamic Representation

Considerable simplification of the dynamical equations can be achieved if suitable averaged quantities can be defined and fluid-like equations used to determine their behaviour. The dynamical relations between the bulk or averaged quantities may be derived by averaging of the equations governing single particle motion which were discussed in the previous chapter. Perhaps the most widely used such simplification is magnetohydrodynamics (MHD). In this approximation, a set of dynamical equations is developed which govern macroscopic behaviour. Quantities may occur in these equations which are derived from a particle treatment. One such quantity is the conductivity, as described below. It allows a relation to be established between the macroscopic electric current and the electric and magnetic fields, but must itself be calculated from a more general theory than MHD.

c. Conductivity, Electric Field, Electric Current

The effect of an electric field applied to a conducting medium is to cause electric currents to flow in that medium. In an isotropic medium with mobile electrons, their acceleration opposite to the direction of the applied field is the primary source of current. The limiting factor on the magnitude of the current is collisions, whose frequency may be denoted ν. In the noncollisional regime of the magnetosphere these considerations are not applicable. However, in the denser ionosphere collisions are important and the effect of electric fields may be considered in a similar manner to that in, say, a metallic conductor. Due to the effects of the magnetic induction field through the Lorentz force, the medium is nonisotropic in its response to the applied electric field. The magnetic effect enters the conductivity expressions which follow in the form of the gyrofrequency $\omega_{cs} = \frac{q_s B}{m_s}$, where the subscript s denotes the species, either i for ions or e for electrons. This electric current due to the applied electric field **E** may be written as $J = \sigma \cdot E$ where σ is the conductivity tensor [see, for example, Parks, 1991, p. 283] with structure, in a system aligned with the magnetic field,

$$\sigma = \begin{bmatrix} \sigma_1 & -\sigma_2 & 0 \\ \sigma_2 & \sigma_1 & 0 \\ 0 & 0 & \sigma_0 \end{bmatrix}$$

where

$$\sigma_1 = \frac{nq}{B}\left[\frac{\omega_{ci}\nu}{(\nu^2 + \omega_{ci}^2)} - \frac{\omega_{ce}\nu}{(\nu^2 + \omega_{ce}^2)}\right]$$

is the Pedersen conductivity, effective in the direction of that component of the electric field which is perpendicular to the magnetic field ,

$$\sigma_2 = \frac{nq}{B}\left[\frac{-\omega_{ci}^2}{(\nu^2 + \omega_{ci}^2)} - \frac{\omega_{ce}^2}{(\nu^2 + \omega_{ce}^2)}\right]$$

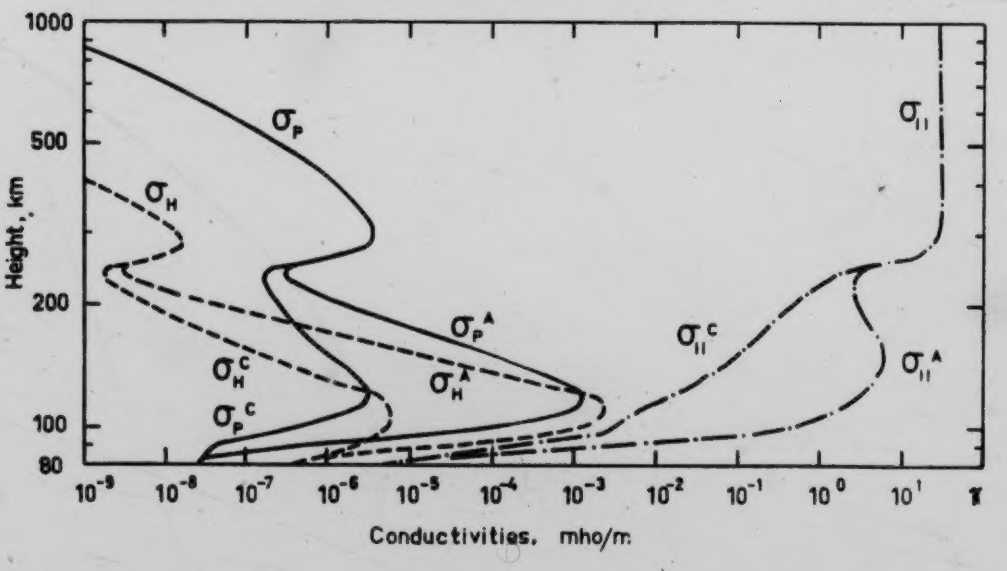
is the Hall conductivity reflecting the magnetic field's deviation of accelerating particles (after a collision) and is effective in the direction perpendicular to both **B** and **E**, and

$$\sigma_0 = \frac{nq}{B}\left[\frac{\omega_{ci}}{\nu} - \frac{\omega_{ce}}{\nu}\right]$$

is the parallel conductivity, unaffected by **B**. Boström [1964] gives forms which are more correct in that the differing collision frequencies between the species (including neutrals) are taken into consideration. However, he notes that at all heights where there is significant conductivity, the Hall and Pedersen conductivities are determined solely by the properties of the ions, resulting in slightly simpler forms for the above expressions. The parallel conductivity has no functional dependence on B since it is cancelled from the denominator by appearing in the numerator in the gyrofrequencies. It essentially depends only on quantities characterizing the electrons, which have great mobility parallel to the magnetic field.

These expressions may be evaluated by substitution of appropriate quantities in and outside auroral arcs, and are shown graphically in Figure 2.1. The conductivity structure has two peaks due to the interplay of number density and collision frequency. The conductivity within an arc is greatly enhanced due to the increase in number density there. This conductivity is strongly peaked at roughly 110 km, which is the height of ionospheric currents used in modelling in this study. Due to the strongly peaked conductivity distribution, the approximation is made in modelling that the currents flow in a thin sheet at 110 km height. Kisabeth [1972] studied the effect of slightly changing this height and concluded that the effects of varying it between 100 and 145 km were not significant, and further that a thin sheet approximation was adequate and that more precise modelling with integration over a layered current distribution was not warranted.

Figure 2.1 Conductivities within and outside an auroral arc. From Boström [1964]. Subscripts H, P, and || denote Hall, Pedersen, and parallel conductivities, respectively. Superscript A indicates results within an arc, C outside. Units of mho/m shown are equivalent to SI units of S/m (siemens per meter).



Since the approximation of a thin conducting sheet is applicable, it is convenient to define height-integrated conductivities Σ and current densities $\mathbf{j}$. The height-integrated conductivity is more properly called the conductance and has SI unit S (siemens). In the case that a two-dimensional Cartesian coordinate system is defined locally having the x axis positive eastward and the y axis positive northward. Ohm's law in the ionosphere may be written [Hughes *et al.*, 1979] as

$$\begin{aligned} j_x &= \Sigma_P E_x + \Sigma_H E_y \\ j_y &= -\Sigma_H E_x + \Sigma_P E_y . \end{aligned}$$

Formulation in a more general coordinate system [Lyons, 1992] but with the desire to separate effects of Hall and Pedersen conductivity, would result in the expressions

$$\mathbf{j}_P = \Sigma_P \mathbf{E}$$

and

$$\mathbf{j}_H = -\Sigma_H \mathbf{E} \times \frac{\mathbf{B}}{\mathrm{B}}$$

which are equivalent to the above with $\mathbf{j} = \mathbf{j}_P + \mathbf{j}_H$. $\mathbf{j}_P$ and $\mathbf{j}_H$ are referred to as (height-integrated) Pedersen and Hall current densities, respectively, with units Am^{-1}.

Indirect observations of the conductivity may be made if the electric current and electric field can be observed simultaneously. This is one motivation for the present approach to measuring currents, since the electric field may be observed remotely by radar or *in situ* by satellites [Kamide, 1988, Section 2.3]. The ratio of Hall to Pedersen conductivities may itself be of interest and can be measured from satellite data under certain assumptions. For example, Sulzbacher *et al.* [1982] used magnetic measurements from the ground and from the TRIAD satellite, along with the assumption that the ionospheric electric field was exactly southward, to deduce that ratio, finding a value which varied from 4 at the northern boundary of the electrojet to about 0.8 near the southern border. In many existing routines for inversion of ground data the conductance plays a crucial role (see Chapter 4). In the method developed in this study, the conductance has no direct role, but could be derived from the modelling results if supplemented by measured electric field.

d. Induction Effects

Faraday's law describes the induction of an electric field as the result of a time-varying magnetic (induction) field: $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$. In space physics, there are two important low-frequency results of this effect in conductors. Within the conductor, currents flow, and the entire set of relevant equations gives rise to wave solutions [Kraus and Carver, 1973, Section 15-10]. These Alfvén waves are similar to the familiar electromagnetic waves but arise through the interplay of electric currents with the plasma (through the $\mathbf{J} \times \mathbf{B}$ force) rather than displacement currents. They are involved with higher frequency phenomena, such as pulsations [Samson, 1991], than are considered here and will not be discussed further. The second induction effect is important, in practice, in regions outside the plasma. This is induction of currents in the Earth or other conductors which may be near a point of observation.

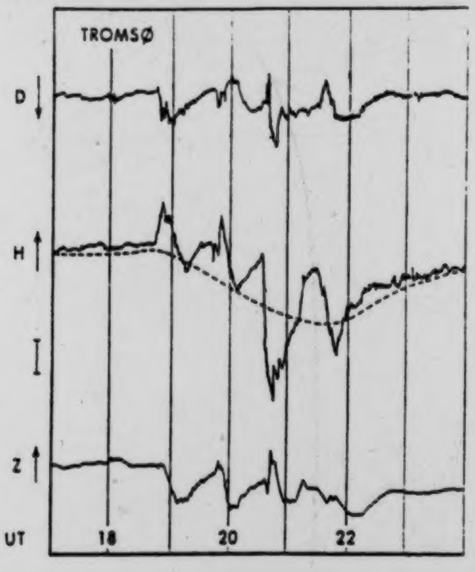
In practice, the induction of currents in the Earth is important due to the proximity of these currents to ground magnetic observatories. In cases of nonuniform ground conductivity, there may be no effective way to incorporate the induction effect, and a nearby station would be considered 'anomalous'. Such cases must be treated individually, and are, in the chapters presenting modelling results. Induction in a uniform Earth also occurs, and is incorporated into the modelling routine as described in Chapter 4.

It has been noted in the previous chapter that induction localized in regions of the magnetosphere may be responsible for particle energization. By Faraday's law, temporal variation of the magnetic field leads to spatial variations in the electric field. If the geometry is appropriate and temporal variation rapid enough, large electric fields and thus forces on charged particles may result. The role of induction is intimately tied, because of the involvement of a time derivative, to the time scale of the event under study. We now examine the relevant time scales for substorms.

e. Substorm Time Scale

The gross physical effects of substorms which are studied here take place on a scale of minutes to hours as shown in Figure 2.2. The short period (high frequency) cutoff is dictated by the desire not to study pulsational phenomena (although they may be of great importance) and by the availability of data with one minute temporal spacing.

Figure 2.2 Prototypical relation of polar magnetic substorms to geomagnetic bay. Major deviations in the H component are spaced by about one hour. Scale bar at left indicates 53 nT in H, 48 nT in D, 72 nT in Z. From Rostoker [1969].

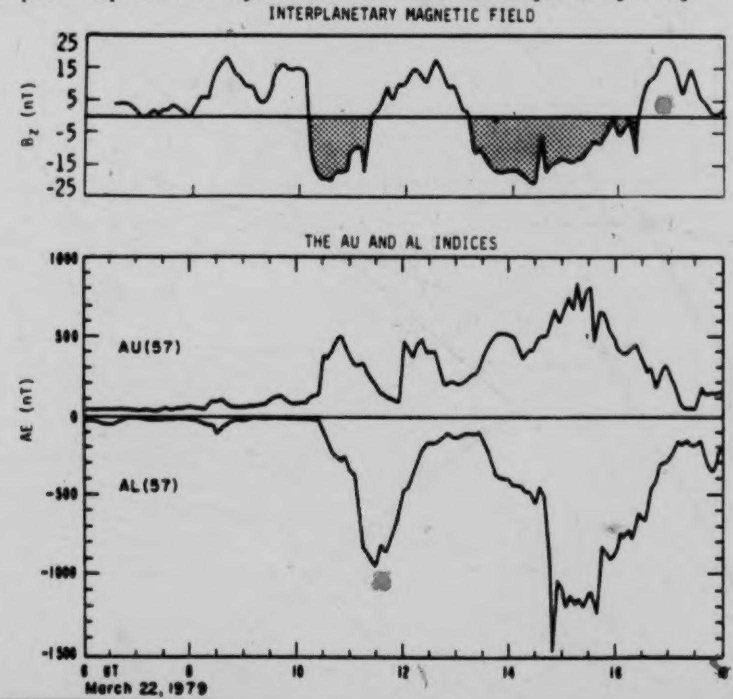


Rostoker [1969] pointed out the existence of two basic time scales in the gross development of auroral current systems. At the time there was controversy about whether the overall layout of such currents was as a one-cell or two-cell equivalent current system. Rostoker identified the one-cell system as being responsible for polar magnetic substorms, with a time scale typically in the tens of minutes. Longer period geomagnetic bay variations were attributed to the two-cell system. Figure 2.2 shows the prototypical case of a sequence of substorms with associated longer-period geomagnetic bay. These data are from Tromsø at corrected geomagnetic latitude of 66.3° N. H perturbations at stations further north (70° to 75°) were found to be consistently negative indicating near-overhead passage of the electrojet associated with the current wedge. Weak H signatures at 19 UT and 20 UT along with negative Z at Tromsø confirm this. Stronger negative H signatures near 21 UT and 22 UT, along with weak and transitional Z, indicate that these disturbances occurred nearer to Tromsø, that is further south than the previous activity. The dashed line indicates the long-period geomagnetic bay associated with this activity. The activity of the two systems is now recognized to be possible simultaneously, with the two-cell system often referred to as the 'driven' system, and the one-cell system as the 'substorm current wedge'. The latter was introduced in Chapter 1 and some of its observational aspects discussed there. As the three-dimensional nature of the auroral zone currents is now firmly established (see introduction to Chapter 5), reference is not usually now made to the cell nature of the equivalent current systems. However, the observation that there are two time scales involved in substorm activity remains valid and is taken to reflect differences in the physical mechanisms or the spatial scale of the regions involved in causing the currents of the two systems.

In ensuing years there has been much discussion of substorm time scales and possible physical implications in terms of overall magnetospheric processes. Discussion is provided by Rostoker [1994, 1995], who points out further that use of certain magnetic indices (in particular AE: see section 3.a.1.b) can obscure the substorm time scales by combining effects from the driven and wedge systems. Further, relating ground magnetic data to satellite or other data can be difficult if use of the index causes timing discrepancies. If activation of the driven system is rapid enough it may be confused with an expansive phase onset if only an index is used for onset timing. In fact, driven system activity may involve near steady-state dissipation of energy in near-Earth regions while there is 'storage' of energy in the tail as discussed in the first chapter. The expansive phase onset would then correspond to 'unloading' of stored tail energy in a not completely understood manner, triggered by an unknown mechanism. In this sense the driven system activity reflects an aspect of the substorm process which may be most readily distinguished during the 'growth' phase.

The triggering mechanism for substorm expansive phases remains in dispute and indeed may be nonunique as are many effects in the domain of chaotic dynamics. Figure 2.3 illustrates the relationship, for the CDAW 6 March 22 1979 substorms[1], between IMF B_z and AU and AL indices computed using 57 stations rather than the normal 12. The use of many stations should in principle allow better timing of expansive phase onsets. In this case driven system activity is seen from 1030 UT on as manifested by a decrease in the AL index and an increase in AU. Shortly after 1100 there appears to be an onset as shown by a steep drop in AL and a period during which AU decreases. Again at approximately 1330 UT this pattern repeats on a larger scale. The general relationship of substorms to southward turning of the IMF is clear; in both cases the driven system activity commences, after a short delay, in synchrony with a southward turning. The expansive phase onsets, however, appear to be more closely related to northward changes in the IMF. This is particularly clear in the case of the very sharp and strong onset at 1450 UT, which follows an impulsive decrease in the magnitude of southward GSM B_z which appears to mark the start of a fairly gradual northward turning. Extension of this prototypical case to general behaviour appears to be unwarranted: northward turnings may trigger some onsets but do not appear to trigger all. This question will be considered in detail in what follows.

Figure 2.3 Comparison of IMF B_z and Auroral Electrojet indices. The relationship of IMF southward B_z to activity is clear, and northward variations in the IMF appear to trigger expansive phase activity. From Kamide and Baumjohann [1985].



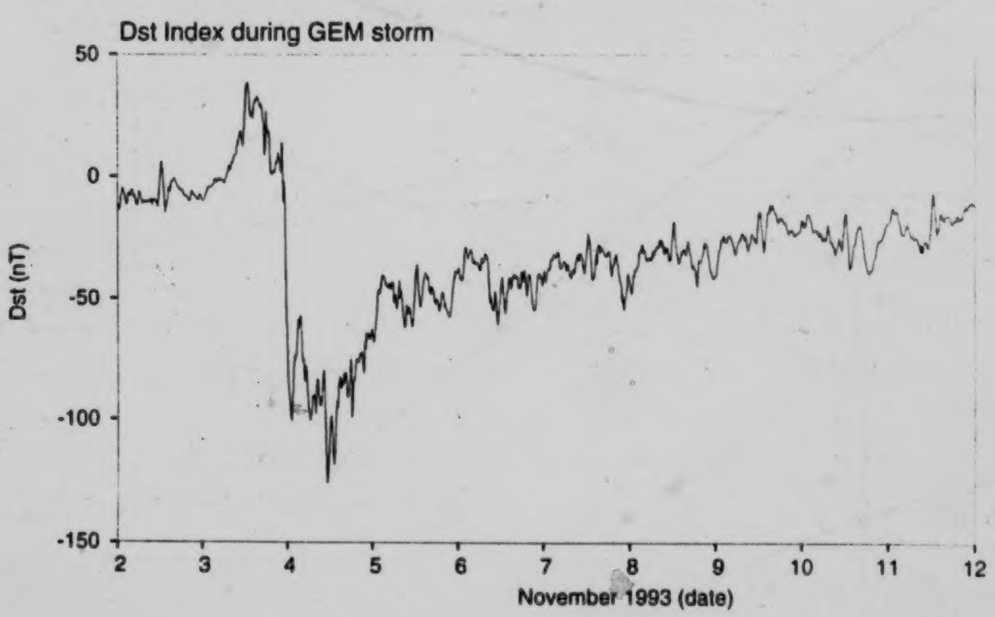
---

[1] This substorm sequence is often regarded as prototypical and has been described extensively, notably in a special section of the Journal of Geophysical Research (Volume 90, Number A2, February 1, 1985, pages 1175-1374).

In a prefatory article about the CDAW 6A substorm, McPherron and Manka [1985] emphasize that the major question related to substorms is the relative importance of energy dissipation directly driven by the solar wind and that driven by release of energy stored in the geomagnetic tail. This question remains largely unanswered. It suggests that attempts to model magnetic activity should include current systems analogous to the driven system and to the substorm current wedge, so that the behaviour of each system may be examined separately.

Another relevant substorm-related timescale is readily observed in the horizontal component at mid to low latitudes. The D_{st} index, derived from low-latitude but not equatorial stations, is generally depressed during magnetic storms for periods of up to several days and this has been shown to be due to the growth of the Earth's ring current [Nishida, 1978, Chapter 4]. A similar but shorter-lived phenomenon occurs during substorms, possibly giving rise to 'partial' or 'asymmetric' ring currents not extending fully around the Earth [Nishida, 1978, p. 127]. Certainly cases are known in which particles, after a substorm injection, display 'echoes' [Lopez and Baker, 1993], so that it is not necessary to consider that injections form only partial ring currents, but rather they contribute to the global ring current [Liu and Rostoker, 1995]. Figure 2.4 illustrates the behaviour of the D_{st} index during an active period, that of the GEM study in early November 1993. One particular onset during this period is treated in detail in Chapter 10.

Figure 2.4 Behaviour of D_{st} Index during a period of storm activity. Initial phase with positive values is followed by the main phase with negative values. Data are from November 1993 and an event from this period is discussed in Chapter 10.



The injection of particles during substorms was discussed in Chapter 1. Such injections can contribute to the ring current as the energized electrons drift eastward and the accompanying ions drift westward, their differential motion causing a current. The initial energization of the bulk of ring current particles may be due not solely to injections, but also partly to non-adiabatic processes occurring during compression of the magnetosphere at the time of storm onset. In any case, D_{st} may be seen in the figure to decline rapidly after an initial phase due to magnetospheric compression. The continued (for hours or days) low values of D_{st} are diagnostic of a storm: the decay of the ring current appears to be a slow process in the essentially collisionless plasma of the near magnetosphere. Impulsive decreases seen in D_{st} may be due to later injections associated with substorms during the storm, which revitalize the ring current. From a modelling point of view, the worldwide depression of the horizontal component must be included in the forward model if mid or low latitude stations are included. However, the generally small D_{st} values during a typical substorm episode are easily overwhelmed by noise at auroral zone stations. Attention must be paid to incorporation of lower latitude stations if a D_{st}-like parameter is to be extracted from the modelling. In global modelling, a D_{st}-like parameter can be determined, but in regional (auroral zone) modelling it should be supplied as an input parameter or may be neglected. These issues are further discussed in Chapter 5[2].

An important consideration in discussion of substorm time scales is the applicable physics, briefly alluded to above. Although Gauss' law $\nabla \cdot \mathbf{B} = 0$ is universally true and has no implications for time changes of the magnetic induction field **B**, Ampère's law in the general case [Slater and Frank, 1949, p. 85] is $\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}$ where **H** is the magnetic intensity, **D** the electric displacement, and **J** the electric current density (A m^{-2}). In a plasma, bound charges are unimportant and the constitutive equation relating **H** and **B** is simply $\mathbf{B} = \mu_0 \mathbf{H}$, where μ_0 is the vacuum permeability ($4\pi \times 10^{-7}$ H/m). However, free charges do move in response to applied fields, giving rise to polarization currents and an effective dielectric constant (in fact dielectric tensor) ε so that the electric field and electric displacement are related by $\mathbf{D} = \varepsilon \mathbf{E}$. If time variations are fast enough, the polarization currents, even in the usual case of a near-neutral bulk plasma, may be large enough that the electric displacement term becomes large compared with the local current density. In this case the situation is analogous to that prevailing in free space, where displacement currents exist although there is no actual current. That situation allows electromagnetic radiation (waves) to occur. In the situation where polarization currents become important in a plasma, wave phenomena may similarly occur. Although an entire rich literature exists about plasma waves, and although they are of importance in studying the magnetosphere [Samson, 1991], from the point of view of modelling in this study it is desirable to use a simpler form of Ampère's law in which time variations may be neglected and in which wave phenomena are not encompassed. As discussed in Section 4.a.1, in this case there is a unique relationship between currents in all of space and the magnetic induction at any given point, given by the Biot-Savart integral. The boundary between magnetostatics and magnetohydrodynamics in the magnetosphere is generally taken to be at periods of tens of seconds. Since wave phenomena often are periodic, it is also the case that averaging over wave periods will effectively remove them from the data considered. For the most part, the work done here is done with one-minute averaged magnetic induction values. It will thus be considered that the Biot-Savart integral effectively relates the observed values to the currents (everywhere in space) which generated them.

---

[2] Using low latitude perturbations, it is also possible to define an asymmetry parameter A_{sym} [McPherron, 1995], which might be of use in separating intensified tail current effects from those of the ring current. A forward modelling study to do this would likely be a feasible project.

When magnetic data with fast time sampling (i.e. at intervals of less than one minute) are discussed here, it will generally be in the context of the use of pulsations as signatures of characteristic events. The use of the damped low-frequency (5-15 mHz) Pi 2 pulsations as signatures of substorm onset is notable in this regard (again see Samson [1991], and also Harrold [1986]). Since data with one minute or larger sampling are used in the modelling, and since stations are generally needed in as many locations as possible, pulsation data, when available, were time averaged to remove pulsations and used along with other data.

We now proceed to a discussion of what types of data are in fact available, how they are obtained, and the coordinate systems which are relevant to their processing.

Chapter 3. Data Availability and Preparation

In this chapter, types of data available for analysis are briefly discussed, with emphasis on the preparation of those actually used in this study. The primary thrust of this work is the use of an optimization technique as applied to ground magnetic data. This optimization technique can in fact be used on any type of data which is able to be represented in digital form and for which a model with variable parameters may be put forward.

a. Fields - Remote and In-Situ

The electromagnetic effects of space plasmas may be measured remotely, primarily from the surface of the Earth, or *in situ*, primarily by satellites. The magnetic field as measured at the surface includes a multiplicity of sources including those from near-Earth space. The electric field at low frequencies as measured there does not include significant contributions from fields in space. Results of satellite measurements of electric fields have been discussed in the introductory chapter, and its role in other modelling techniques will be discussed in Chapter 4. The modelling technique used in this study is based on the magnetic field and does not directly involve the electric field. We proceed by discussing the magnetic field measurement and preparation techniques in detail.

1. Magnetic Fields

Magnetic fields are produced by moving charges or by changing electric fields. As such, they carry information about charge motion (usually differential motion of charges to form electric currents) and about electric fields in a non-static situation. Since they may be detected at great distances from a source, they convey information about distant conditions. This is both an advantage and a drawback in that many sources may contribute to an observed total magnetic field, and it is in practice difficult to disentangle their signals. In the forward modelling approach one attempts to do so by proposing a certain behaviour from known source configurations and matching the data by varying associated parameters.

Magnetic fields also exert pressure, affect charged particle motion, and store energy. These aspects play a role in determining the structure and dynamic behaviour of near-Earth space, as we have seen in previous chapters.

a. Magnetometers

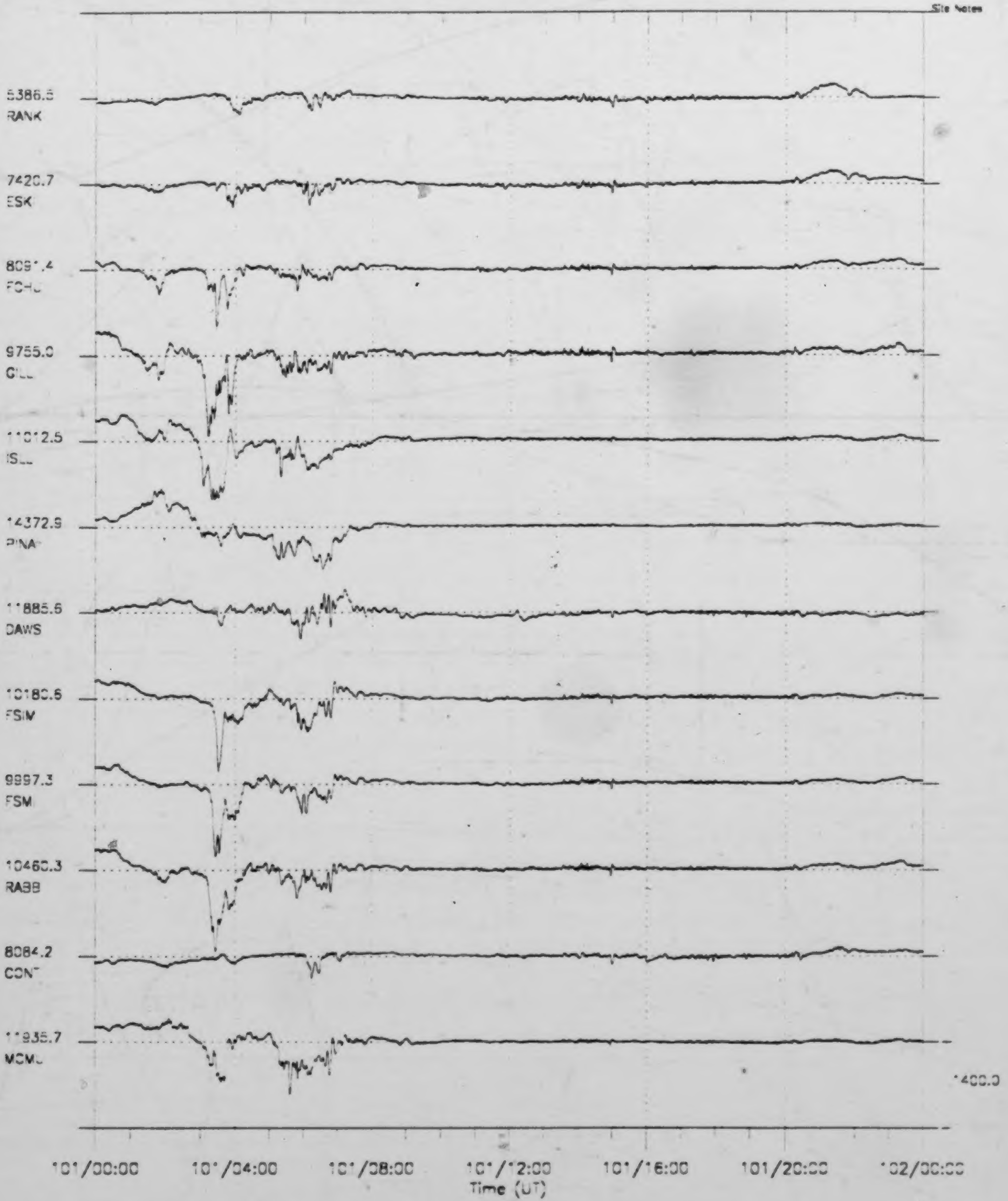
The history of measurement of the Earth's magnetic field in its many aspects is discussed in detail in the monograph by Chapman and Bartels [1940]. The ancient Chinese may be regarded as the first to acquire knowledge of the Earth's field, to the extent that they may have known of magnetic declination [Temple, 1986]. Technical progress in magnetometer design has been substantial since then, and modern systems for measurement of vector magnetic fields to subnanotesla precision are generally based on fluxgate principles [Vacquier, 1993; Forbes, 1987]. In some other types, the total field is determined very accurately while individual components cannot be measured. These are referred to as scalar magnetometers. In this work it may be assumed that the measurements were taken with fluxgate magnetometers and that the components have been measured to at least 0.1 nT accuracy. Thus, accuracy of the magnetometers is not usually a consideration in error analysis here. Frequently, in fact, data are only stored (and thus available for analysis) to the nearest nanotesla in any case.

b. Ground-Based Magnetometers

Already in the middle of the 19th century, many observatories operated worldwide, largely due to the efforts of Carl Friedrich Gauss (1777-1855). Part of Gauss' motivation in promoting a worldwide observatory network was his mathematical techniques which allowed the separation of an observed field into parts due to internal and to external sources. Given sufficient spatial coverage the origins of the Earth's magnetic field could thus be hoped to be established. Networks of observatories are also very useful in studying the localized and variable fields associated with auroral activity. Figure 3.1 shows data from much of western Canada obtained by the CANOPUS array. Since that time developments in the understanding of the field have been numerous: however it is still the case that further progress is impeded by the sparsity of magnetic observatories. Gauss' goal of understanding the origins of the Earth's main field also has not yet been completely achieved.

In addition to the main field which makes up by far the largest contribution, ground-based magnetometers detect fields due to both external sources and sources within the Earth which are not considered to be parts of its main field. The main field is considered to be of deep origin, has significant spherical harmonic power only up to a certain finite degree and order (that is to say, is spatially fairly smooth), and varies slowly over time. External sources typically have temporal frequencies higher than those of the main field, up to several Hz. Even in polar regions where they may have a large effect on compasses, their magnitude is much smaller than that of the main field. The signals studied here are due to external sources, known to be located in the ionosphere and lower magnetosphere. These are not, however, the only sources of external field in the frequency range considered: the other sources must be considered noise and removed from the data. Sources within the Earth may be essentially DC signals, due to crustal magnetization or local field sources, or may be induced fields arising as a result of variations in the external fields. Local DC signal is generally considered to be an 'anomaly', and may be removed (being DC) by subtracting 'quiet day' signals observed at the same place. Often, however, the presence of a DC anomaly is an indicator of complex structure in the Earth's crust near the station, and these stations also respond anomalously to variations in the field. Such response may aid in modelling the local geological structure, but is inimical to study of processes arising in space. It cannot in practice form part of a forward modelling scheme and here anomalous stations must be treated by a reduced weighting in analysis. Such cases are described where relevant in the analysis chapters which follow. Temporal variations can also occur which are not of interest in studying the currents associated with the aurora. These are now briefly described.

Figure 3.1 Magnetograms from the CANOPUS array for April 11, 1997[1]. The northward component (X) of the magnetic field is shown as a function of UT.



---

[1] Courtesy of Canadian Space Agency at http://www.dan.sp-agency.ca/www/canopus_home.html

An important source of external signal is the local atmosphere, particularly in daytime. This is primarily manifest in the form of the 'solar quiet' variation, S_q. In response to solar tides and heating, the conducting ionosphere has vertical motion, which induces an electric field since to a good approximation $E = -v \times B$. This field, applied to conducting layers of the atmosphere, in turn causes current flow. This current-carrying layer has been directly observed by sounding rockets although such measurements are few and only of the scalar field due largely to the great difficulty in obtaining precision vector measurements given the poor pointing accuracy possible on a rocket. Observations at midlatitudes by Yabuzaki and Ogawa [1974] showed the current layer to be at 101 to 110 km altitude on a quiet day evening at Kagoshima (31.25° N geographic). Well before the explicit demonstration that an ionospheric current system caused S_q, its equivalent current system was deduced. Now we know that this equivalent current system is in fact a physical current system. Figure 3.2 illustrates the global nature of this current system near equinox, with 25 kA of current flowing between adjacent contour lines. The current is most intense in equatorial regions due to the horizontal orientation of the main field there and greater tidal and heating effects on the atmosphere. The system is approximately centred on noon local solar time; at this equinoctial time it is also centred on the equator. At other times of the year the overall system moves latitudinally with the Sun: further north in summer and further south in winter (Northern Hemisphere seasons). As observed from the ground the S_q variation has latitude dependence and local time dependence as would be expected from the location of the current systems. Figure 3.3 shows the behaviour of averaged S_q at various latitudes. Tashkent, at 69.3° E, 41.4° N (geographic coordinates), is illustrative. As the day progresses, it is carried to larger local times, first passing under a region of essentially southward ionospheric current flow which produces a ground magnetic perturbation which is mainly eastward, or in the sense of positive Y. Thus the Y trace here shows a prenoon maximum. At local noon the overhead current flow is essentially westward producing the noon minimum in the X or northward magnetic component. For postnoon local time the overhead current flow is essentially northward and a Y minimum is seen. The X or northward component is maximal when east-west current is overhead, which is at local noon, and varies in amplitude with the latitude of the station. Stations at approximately 30° latitude show little variation in the north-south component as east-west currents do not pass overhead. These stations do see ample Y or eastward component as they rotate under the Sun. Depending on the hemisphere, first one sign and then the other of north-south current flow passes overhead. Since the overall half-width of the system, as seen in Figure 3.2, is about 90° of longitude or 6 hours of local time, the Y component extrema are separated by about this amount of time. Figure 3.4 shows total Y magnetic variation at Fredericksburg, geographic latitude 40° N, during the UT day of April 3, 1986 (see chapter 6 for a thorough analysis of part of this day) and on another April day. Local noon at this station is at roughly 17 UT and Y component variations peak at roughly 14 UT and have minima at roughly 19 UT. This behaviour is consistent with that described above, and the figure may be compared to the Y component curve at Tashkent in Figure 3.3 given the comparable latitude. Also of note is that the April 3 data show large perturbations at 0 to 8 UT which are not present on the quieter day. These occur in local night, and the quiet day shows that S_q effects, as expected, are minimal during this time period. However, while the perturbations in question are large compared to what is usually seen, they are of approximately the same size as the S_q variation. If dayside data are used, then, it is clear that S_q effects must be allowed for in some fashion. This is usually done by simple subtraction from the studied perturbations of those of a quiet day not more than a week or so from the day being studied. Since S_q varies with season, a quiet day too far removed in time from the day studied would introduce spurious S_q effects or not sufficiently remove those present. This technique assumes that the S_q variation is not much affected by the processes responsible for the disturbance being studied. In practice, this method of reference to a quiet day has been used in most data preparation in this study. Another method for S_q removal is to grid S_q as a function of known variables such as day of year, location, and local time and remove this function appropriately evaluated. This approach removes, by averaging, random effects which may be present on any given quiet day. Once more the implicit assumption is that the quiet day variation removed is appropriate under the conditions studied.

Figure 3.2 S_q ionospheric current system under equinoctial conditions. 25 kA flows between each adjacent pair of contour lines, with direction of circulation shown by an arrow in each hemisphere. The system, fixed in local time, causes varying perturbations as a station passes under it with Earth rotation. From Parkinson [1983].

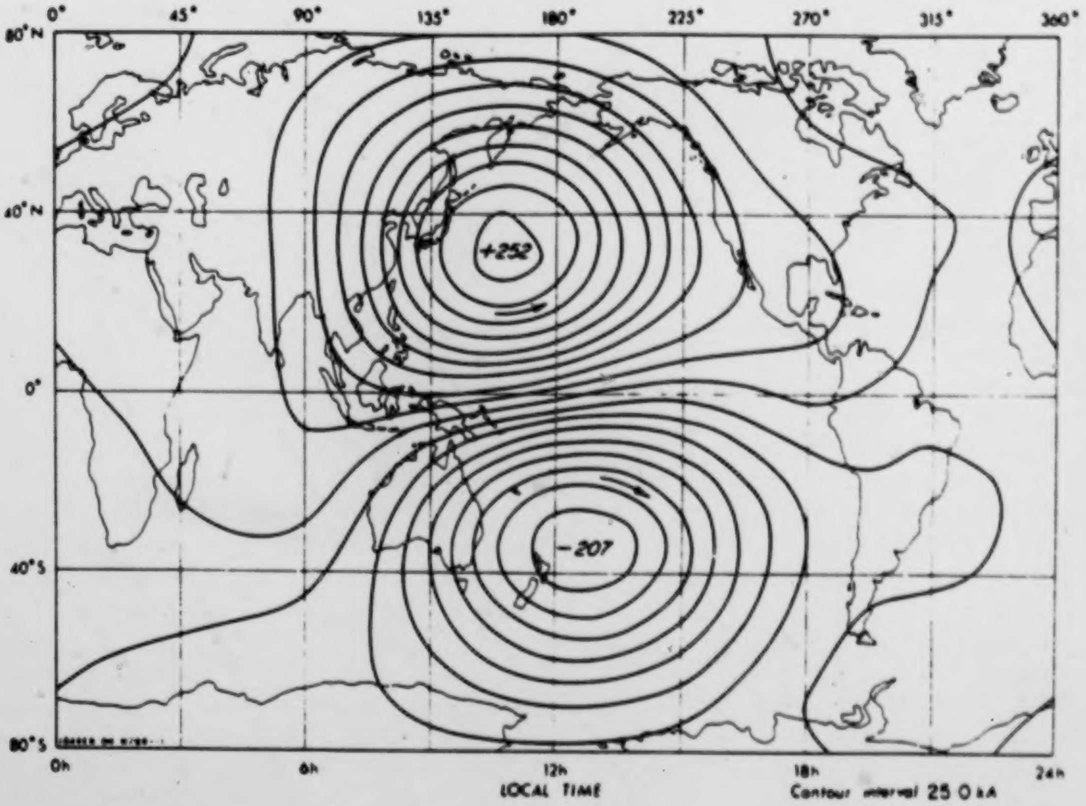


Figure 3.3 Averaged quiet day variations at various latitudes, showing principally S_q variations. Local time is indicated under each stacked component as M (midnight) or N (noon). From Parkinson [1983].

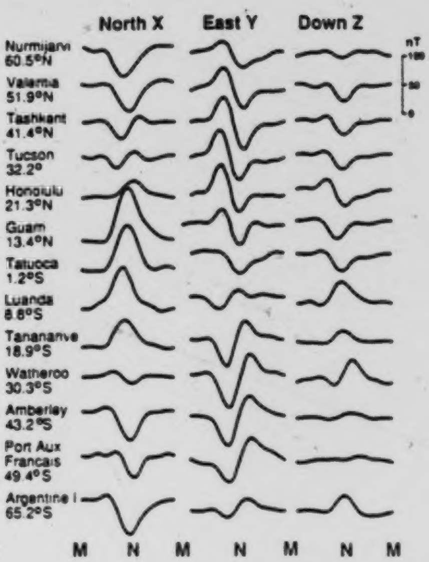
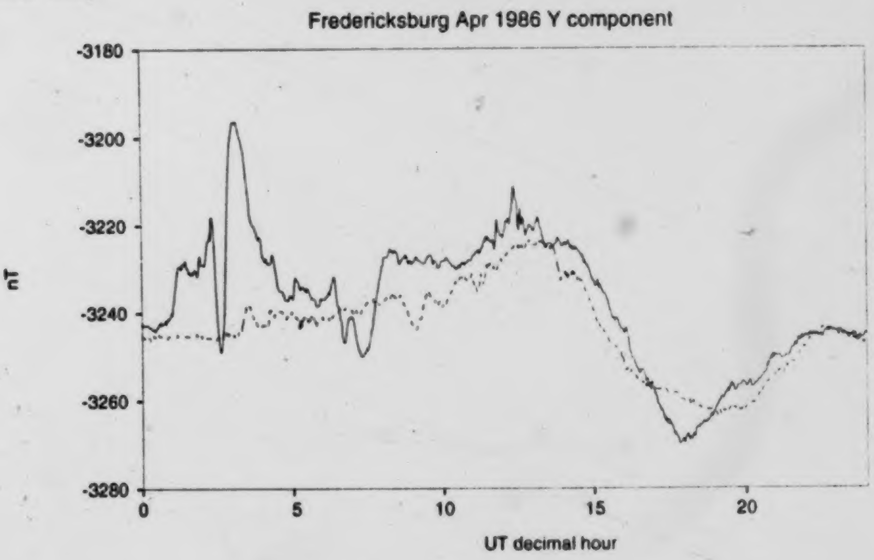
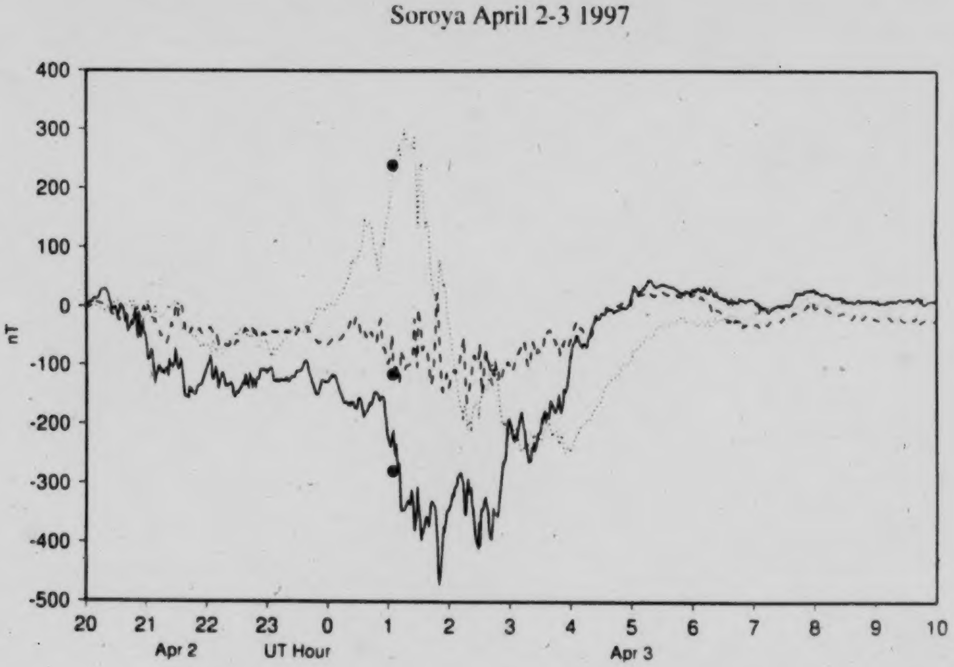


Figure 3.4 Perturbations observed at Fredericksburg (77.5° W, 38.3° N geographic), showing total Y component on April 3, 1986 (solid line) and on a quieter April day (dashed line).



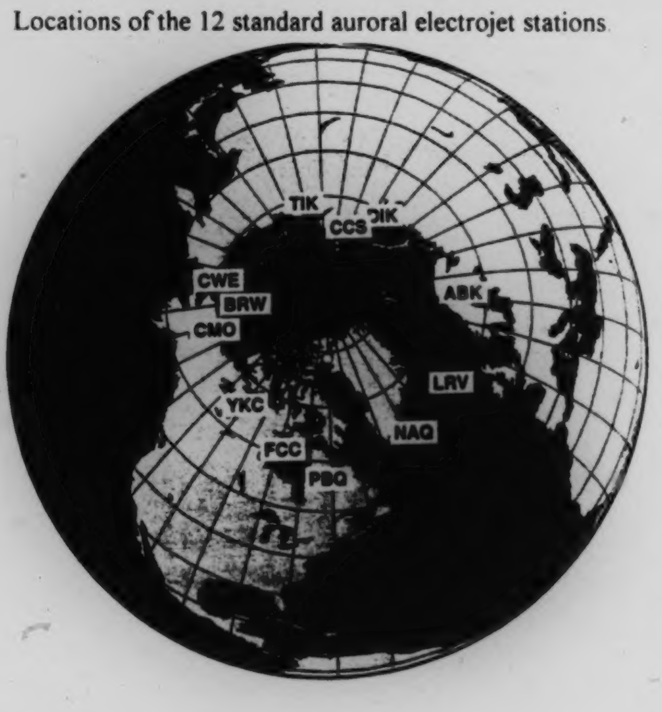
It is implied above that nightside data do not need to be corrected for S_q, although this might depend to a certain extent on the season. In some cases it is most practical to work with only nightside data and to choose a quiet interval as a reference point. There may be little choice, as campaign or study interval data such as those of CDAW may not include appropriate quiet intervals in the data set. Figure 3.5 justifies this approach by showing data from Soroya in Northern Scandinavia (22.8° E, 70.6° N geographic) taken from the CDAW 9 data set and referenced to a baseline value from 20 UT April 2 1986. At 1:05 UT on April 3 1986 each data line has its corresponding value, as corrected by comparison to a quiet day (April 4, 1986) and averaged over 5 minutes, shown by a filled circle [B. Emery, personal communication, 1995]. It is seem that in this case simple removal of a reference value during a quiet interval on the same night has produced an accurate representation of the perturbations. As this station is roughly one hour of local time ahead of Greenwich, night hours are from about 17 UT to 5 UT. S_q effects should be mainly near local noon (11 UT, just off the diagram) in the north-south component at this latitude (compare figure 3.3) but are not easily seen in the data shown. They are typically small, while substorm-related variations are large, at auroral zone stations, particularly on the nightside. However, the method used in this study attempts to include global stations, including those at middle and low latitudes. For some of these, particularly on the dayside, S_q removal by one of the methods described is important.

Figure 3.5 Comparison of magnetic perturbations deduced from two methods of correcting night time variations. Lines indicate the H, D, and Z components (solid, dashed, and dotted, respectively) for April 2-3 1986, as referred to 20 UT April 2. Filled circles indicate data points at 1:05 UT as corrected by reference to the April 4 1986 quiet day. The two methods have produced comparable results at 1:05 UT



Since the magnetic field is present throughout three-dimensional space and variable in time, visualizing it is difficult. Representation in various graphical forms, notably as stacked magnetograms as in Figure 3.1, may be useful (other forms were explored by Kisabeth [1972]). Modelling, either through purely mathematical expressions, or through parameters having some physical meaning, may also be helpful simply to aid in data display. Another method to facilitate use of ground data from many locations is though an appropriate magnetic index. The use of indices, such as D_{st} referred earlier, has been widespread. For auroral zone perturbations, the auroral electrojet (AE) index has been the most widely used single number giving an indication of global activity. AE is based on superposition of H perturbations from many observatories. The index is produced as a difference of outer envelopes of the most positive and most negative perturbations from all of the observatories involved. The outer envelope of the most positive perturbations is referred to as AU (U for 'upper') and that of the most negative as AL (L for 'lower'). In active times AL is often strongly negative and dominates, so that AE in such cases is comparable in magnitude to AL, but positive. Clearly, since the auroral electrojet indices are outer envelopes only, much information is lost even from the H component in forming the index. Also, the other components are not incorporated into the index. What has been discussed here can be applied to produce indices from any set of stations. In practice, 12-station auroral electrojet indices based on certain stations spaced roughly 2 hours of local time apart, and all at roughly the same magnetic latitude (65°), are widely used. Unless otherwise specified, auroral electrojet indices are taken to be based on the above procedure applied to the standard stations, whose locations are shown in Figure 3.6.

Figure 3.6 Locations of the 12 standard auroral electrojet stations.



c. Spacecraft Magnetometers

The advent of Earth-orbiting satellites in 1957 opened the possibility for *in situ* investigation of the magnetic field in near-Earth space. Initial investigations by the Explorer 1 and 3 satellites in 1958 used the effect of the magnetic field on trapped high-energy particles to explore its geometry [Burrows, 1990, pp. 79-80], although direct field measurements were not done by these spacecraft. The first spacecraft magnetometer flew in 1958 and by 1961 vector measurements were being taken [Parks, 1991, Chapter 3]. Very soon it was realized that the field in near-Earth space was not completely dipolar: 1961 measurements from Explorer XII led to the discovery of the magnetopause. Slightly later, Ness and coworkers [see again Parks, 1991] observed abrupt reversals in magnetic field direction best explained by the presence of a tail current sheet. Large amounts of spacecraft data have been gathered in intervening years and have contributed to our view of the large-scale structure of the magnetosphere, especially when interpreted in conjunction with other data sets. Some of these data, from magnetospheric spacecraft on distant orbits, such as ISEE (recall Figures 1.15 and 1.16), have already been briefly discussed. An example of behaviour at synchronous orbit has also been shown (Figure 1.20) and the continuous presence of satellites at this location for Earth monitoring has produced important data sets. In conjunction with surface modelling, however, the low Earth orbit class of spacecraft is important and data from this class (typified by Triad, already discussed above) will receive the most attention here.

Recent low Earth orbiting spacecraft carrying magnetometers have included Viking and DMSP F7. Viking operated line-of-sight to Sweden and was generally at about 2 R_E (13500 km) altitude when producing images. Although calculation of the magnetic footprint is straightforward, this high altitude is not favorable for examination of small-scale structure of field-aligned currents in the near-Earth region. This region is better explored by low altitude spacecraft such as DMSP F7 which may nominally be considered to make polar passes at about 800 km altitude (similar to Triad) and thus to respond mainly to field-aligned currents near the ionosphere [Lopez *et al.*, 1991] and somewhat to ionospheric currents.

The satellites of the Defense Meteorological Satellite Program (DMSP) are a series of low-altitude (835 km), polar-orbiting (98.7 degree inclination) satellites. The orbital period of 101 minutes makes the satellites Sun-synchronous. To a large degree, the magnetosphere is also aligned with the Sun-Earth line and a Sun-synchronous satellite samples an equivalent region on each orbit. There are normally two DMSP satellites operating at a time; one in a 0600-1800 local time (LT) orbit and one in a 1030-2230 LT orbit. During the CDAW events modelled in subsequent chapters, DMSP F7 was in a 1030 LT orbit. Its companion F6 was in a 0600 LT orbit but did not carry a magnetometer. The primary purpose of the DMSP satellites, as military weather satellites, is to observe the tropospheric weather using a 2.8 km resolution visible and infrared imaging system. Incidental to this primary mission is the observation of aurora under suitable activity and lighting conditions (see section 3.c.3).

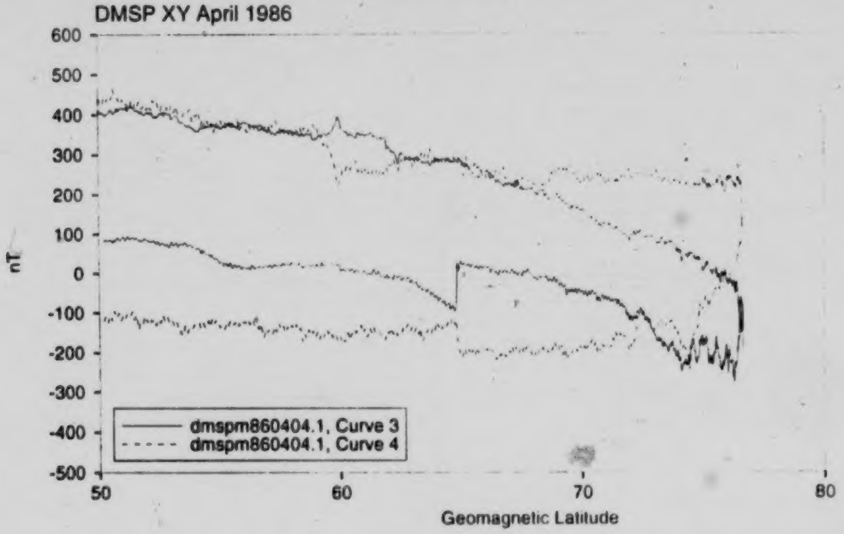
Magnetometer data as supplied [F. Rich, personal communication, 1994] are in ASCII form in logical groups of one minute. The data are almost continuous, with occasional gaps of a few seconds to a few minutes due to telemetry errors. There are a few gaps of hours due to malfunctions at tracking stations, corrupted raw data tapes, and miscellaneous problems. Ephemeris information was supplied for the start of each one minute of data and intermediate positions were calculated as a function of time using splines [Press *et al.*, 1992, Section 3.3]. The ephemeris gives the date and time at the beginning of the minute, the position of the satellite (in fact the footprint traced down magnetic field lines) at this time in geographic and corrected geomagnetic coordinates, the model magnetic field vector (nominally IGRF 80, but in data used here a provisional IGRF 1983.5 was used [F. Rich, personal communication, 1994]) in spacecraft coordinates, and the number of seconds of data in the associated data file. The data file has up to 60 records (or lines) per minute. Each record gives the second of the minute and the 1-sec average of the difference vector (measured magnetic field minus model field). Invalid data were marked by subtracting 1.E6 from the sample. The raw data contained 20 vectors per second. Data supplied for this study were from the standard database and included data from only magnetic latitudes greater than approximately 48 degrees (although the raw data base contains data from all latitudes). Data were supplied in a coordinate system with one component parallel to the spacecraft velocity vector, and for use here required rotation about the vertical into an Earth-fixed coordinate system.

The spacecraft coordinate system has:
+X downward along the local vertical;
+Y perpendicular to +X, in the orbit plane roughly parallel to the velocity vector; and
+Z perpendicular to +X and +Y and parallel to the orbit plane normal
Transformations were performed from this coordinate system into Earth-centered systems using ephemeris data. The spacecraft velocity vector must be accurately determined for such transformations. It was found that interpolation in the one-minute ephemeric table produced uncertainties in the pointing direction with period of about one minute. Such uncertainties in turn produce spurious variations in the transformed magnetic data which are comparable in effective spatial extent to those associated with the auroral zones. Thus a rough orbital fit based on the extremal latitude attained on any given orbit was preferred. An effective orbital inclination derived from this allows a fit with very smooth variation of the spacecraft pointing angle and no spurious effects due to inclination.

The body-mounted (as opposed to boom-mounted) fluxgate magnetometer flown on the DMSP/F7 satellite was a proof-of-concept instrument. The magnetometer returned data on the near-Earth field from December 1983 to October 1987. Numerous spacecraft background signals are present in the data, as will be seen and discussed in data analysis chapters. The magnetometer has been evaluated as almost, but not quite, good enough for modelling the main geomagnetic field [Ridgway *et al.*, 1989]. In practice, use of the data in an absolute sense is not necessary in this study, and the near-repeatability of the orbits favors an approach in which a quiet day is used as a baseline, rather than use of an absolute baseline based on the subtraction of the main field appropriate to the spacecraft location. The subtraction approach was attempted, using the appropriate IGRF model, but did not yield satisfactory results. Such baseline problems are not restricted to this study: DMSP data has been little used (perhaps underused) due to them, and recently an attempt has been made to obtain better magnetic data from this satellite series [F. Rich, personal communication, 1995].

Figure 3.7 illustrates artifacts in the DMSP/F7 magnetometer data during a northern hemisphere pass on April 4 1986, which was geomagnetically quiet in the early UT hours during which the pass occurred. No strong features attributable to natural currents are evident in this plot. The strongest features are artifacts. On the bottom half of the curves, at approximately 65° geomagnetic latitude, is a level shift likely due to tape recorder activity (most but not all of these are removed in processing). On the top (nightside) portions of the curves at 60° to 62° is a signature clearly associated with shadow entry and visible on all passes examined at this local time. Apart from the likely shadow entry signature, the changing of background levels with position of the spacecraft is not explained here: rather the levels as presented are taken as quiet time repeatable levels and used as a basis of comparison (i.e. baseline) when these data are analysed. Data used in this manner allow determination of the locations and strengths of current systems; further, the perturbations may be correlated with particle data available from the same spacecraft at the same time. That is not done in this study, which is focussed on magnetic data.

Figure 3.7 Geographic X (solid) and Y (dashed) components observed by the DMSP/F7 magnetometer on April 4 1986 on an early UT northern hemisphere pass. Bottom portion of each curve corresponds to dayside.



As alluded to in Chapter 1, the magnetic field of the solar wind is also an important quantity in the solar-terrestrial interaction. Since 1973 (October 26 launch date) the IMP-8 spacecraft has been the 'workhorse' at measuring properties of the near-Earth solar wind [King, 1982]. During its period of about 12.5 days, the spacecraft is typically in the solar wind about 60% of the time, and is between 25 and 45 R_E from Earth. Its data are sometimes used in magnetospheric studies while it is in the tail. Although the data are now available in near-real time[2], the chance of the spacecraft being in the solar wind is no longer the limiting factor in obtaining its parameters. The acquisition of data from the spacecraft is now sporadic, presumably for operational reasons. More remote from Earth but always in the solar wind, and with data almost constantly received as of this writing (summer 1997), the WIND spacecraft now provides valuable solar wind data[3]. WIND was launched on November 1, 1994 and is at the Earth-Sun L1 Lagrangian point. At typical solar wind speeds this provides one hour 'warning' of active conditions upstream, but in timing studies the temporal offset, far more than that from IMP-8, is a disadvantage. IMP-8 and WIND measure various plasma properties in addition to the magnetic field.

b. Particles

Since auroral emissions are largely due to precipitation of energetic particles into the upper atmosphere, particle detectors on satellites can provide information about aspects of auroral phenomena. Interpretation of particle data can lead to inferences about resulting ionospheric conductivity [Knipp, 1989] and about magnetospheric source regions. The primary purpose of examining particle data in the present study is to assist in locating auroral features and for this the low altitude data from DMSP are most useful. In some cases, particles will be discussed in other contexts. These include detection of injected energetic particles in synchronous orbit as a result of substorm activity, and use of particles as diagnostic indicators of spacecraft residence in certain areas of the magnetosphere.

The following description of DMSP particle detectors is based on descriptive computer files from the National Geophysical Data Center. The DMSP particle detector, SSJ/4, provides a high dynamic range energy spectrum of the low (typically keV) energy particles that cause the aurora and other high latitude phenomena. The SSJ/4 instrument was designed to measure the flux of charged particles as they enter the Earth's upper atmosphere from the near-Earth space environment. It consists of four electrostatic analyzers that record electrons and ions between 30 eV and 30 keV as they flow past the spacecraft toward the Earth. The instruments "look" toward the satellite zenith. The curved plate detectors allow precipitating electrons and ions to enter through an aperture. Electrons and ions of the selected energy are deflected toward the target by an imposed electric field applied across the two plates. The two low energy detectors consist of 10 channels centered at 34, 49, 71, 101, 150, 218, 320, 460, 670, and 960 eV. The high energy detector measures particles in 10 channels centered at 1.0, 1.4, 2.1, 3.0, 4.4, 6.5, 9.5, 14.0, 20.5 and 29.5 keV. There is effective overlap of the highest of the low-energy channels and the lowest of the high energy channels. At high counting rates these detectors are found in practice to produce near-identical counts [F. Rich, personal communication 1994]. Differential particle fluxes (in units of electrons/cm^2/s/steradian/eV) may be as large as 10^{10} at the lowest energies. The detectors also record high energy ions that penetrate both the satellite and the instrument but discussion of particle data here is qualitative and this is not considered further. Each detector dwells at each channel for 0.09 seconds from high energy channel to low. A complete cycle is sampled each second. In this study the particle data were combined with positional data obtained from ephemeris files as described in the section on satellite magnetometers.

---

[2] http://ndads.gsfc.nasa.gov/ndads/archives_form_IMP8.html

[3] http://www-spof.gsfc.nasa.gov/istp/wind/wind.html

Particle data from synchronous orbit is primarily available from the GOES series of spacecraft and from Los Alamos sensors on military spacecraft. The GOES Energetic Particle Sensor (EPS) is part of a standardized package known as Space Environment Monitor (SEM) carried on GOES and other spacecraft since the 1970s [Fritz and Neely, 1982]. The EPS aims mainly to detect high energy particles (protons and alphas) of external origin but some trapped particles are also detected (high energy electrons of 2 MeV or more and the lowest energy proton detector band of 0.6 to 4.2 MeV) [Wilkinson and Ushomirskiy, 1994]. The Los Alamos instrument package is sensitive to electrons and protons in lower energy ranges, which are primarily trapped. Substorm injection of such particles is well known and was referred to in Chapter 1. The instrument is described by Baker *et al.* [1982] and an example of proton injection associated with a substorm onset, and subsequent drift echoes, is presented there.

Figure 3.8 Dispersive injections of protons into the synchronous orbit region. Observations are from the ATS 5 particle detectors. Energy increases downward on the main plot. The AU and AL indices are shown; large AL bays are highly correlated with extrapolated infinite energy injection times. The D_{st} index is also shown.

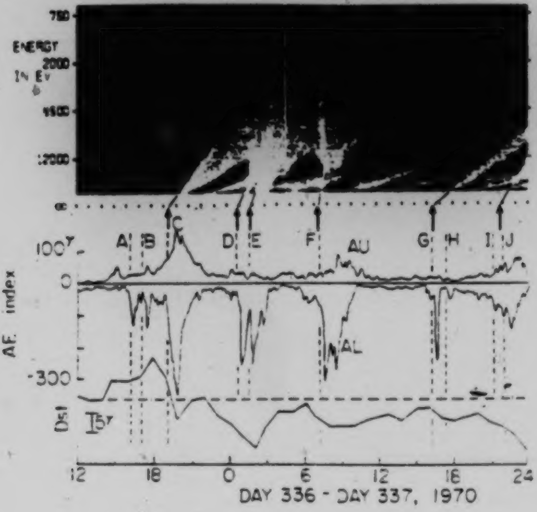


Figure 3.8 illustrates proton injections and their association with substorms as deduced from the AE index and Dst (partially, at least, caused by the injections) in an early study by Kamide and McIlwain [1974]. The injections as observed at the spacecraft (ATS-5) are generally dispersive due to their being associated with midnight sector activity, the differential drift velocity of protons depending on energy, and the fact that the geostationary satellite is not continually near the midnight sector. The one injection with little dispersion is labeled 'F' and this was observed when the spacecraft was very near the midnight sector. In the more general case the spacecraft was not near the point of injection, and the near-linear increase in particle velocity (approximately $15.2\ E\ L^2\ \mathrm{ms}^{-1}$ westward for protons with energy E keV bouncing near-equatorially at radius L Earth radii [Lyons and Williams, 1984, p. 25]) with energy means that the most energetic particles arrive first at the spacecraft, followed by less energetic particles. Extrapolating the curves of arrival versus energy to infinite energy allows approximate determination of the injection times. In several instances in the analysis chapters of the present study, injections as observed by synchronous spacecraft at various longitudes prove useful in localizing and timing substorm onsets.

c. Imaging

Auroral imaging has traditionally been done with ground photography. Photographic triangulation allowed the establishment of the altitude of auroras as being above about 100 km. All sky imaging, in which an optical system produces an image of the entire dome of sky visible from a site at any given time, is useful in establishing the larger scale relationships among auroral features. With several such imagers, the scale may be extended to be nearly global, as exemplified by the pioneering work of Akasofu [1964]. Here, however, few all sky images will be referred to, partly due to the availability of satellite images at times of interest.

Only since the 1970's has it been possible to image the aurora from above using satellite-mounted cameras. A wide, sometimes global, field of view from one instrument is an important advantage of satellite imaging. Correlation between forms may be done without the obstacle of the local horizon. Further, the restriction to observing radiations which pass through the Earth's atmosphere is lifted, although there may still exist absorption or scattering between the emitting layers and the spacecraft.

Another imaging technique which attempts to use the electromagnetic spectrum effectively is meridian scanning photometry [Rostoker *et al.*, 1995]. Although viewing is restricted to one dimension, generally several spectral lines are imaged simultaneously. Auroral behaviour at several wavelengths, through time and along the scan dimension (the meridian) can be effectively displayed. Meridian scanning photometry from the CANOPUS array is used in a demonstration study with CANOPUS data in Chapter 9. It is noted here, but will not be demonstrated, that the fitting technique used for magnetic data in this study could be readily applied to meridian scanning photometer data.

The images most used here are those produced by the Viking spacecraft, which flew in 1986. These have the advantages of good spatial and temporal resolution (since exceeded by the Freja spacecraft). Viking's constraints of narrow field of view and rapid polar passes do not affect this study since events using Viking data were selected for data quality. Images from Dynamics Explorer are also referred to. Like those of Viking, these are in the vacuum ultraviolet spectral region. These images have larger areal coverage but less resolution spatially and temporally than those of Viking. The more recent POLAR spacecraft provides images which are in a sense intermediate to those of Viking and DE, having less resolution in space and time than Viking but more than DE's spin scan camera. We now explore some details of these image sources.

1. Viking

The Viking spacecraft was launched on February 22, 1986 carrying a complement of instruments for investigating auroral and magnetospheric processes. It continued functioning into the late autumn of that year. Spacecraft instruments and initial scientific results were presented in a special section of the April 1987 issue of Geophysical Research Letters [volume 14, pages 379-482].

The imaging experiment, labeled V5, was based upon a pair of all-reflective f/1 cameras [Anger *et al.*, 1987b] with field of view 25°, that ended up slightly misaligned but which basically image the same field in two different vacuum ultraviolet passbands. Camera 0 passed 134 to 180 nm, mainly corresponding to emissions from the LBH bands of molecular nitrogen (N_2) and some from a line at 149.3 nm due to atomic NI. Camera 1 detected from 123.5 to 160 nm, where emission is largely due to neutral atomic oxygen (OI) multiplets at 130.4 and 135.6 nm, and to some of the LBH bands. Camera 1 ceased to function on May 5, 1986 [Vallance Jones *et al.*, 1987]. In principle the UV aurora could be observed on the dayside, but filters which allowed passage of longer wavelengths resulted in low signal to noise ratios in imaging these regions. Due to the ground station being in Kiruna, Sweden, and the use of only real-time, line-of-sight operation, imaging was restricted to the northern hemisphere. The effective exposure of an image is approximately one second and images could in principle be acquired on each spin of about 20 seconds. In practice, operation of several instruments in addition to V5 and bandwidth requirements often limited the imaging to about once per minute (the nominal time unit of this study due to availability of one-minute magnetic data).

UV images from Viking give an indication, on a 20-second to several minute timescale, of auroral luminosity, which in turn indicates regions of particle precipitation and thus current flow [Marklund *et al.*, 1987a]. An electrodynamic inversion method based on Viking images and Viking magnetic/electric data, along with observations from DMSP, was developed by Marklund and coworkers [1987b].

2. Dynamics Explorer

The Dynamics Explorers (DEs) were a pair of polar orbiting spacecraft on coplanar polar orbits [Fujii *et al.*, 1994]. DE 1 was launched in August 1981 and carried an imager. Its initial orbit had apogee above the northern polar cap at 3.63 R_e, enabling hemispheric views to be taken with 12 minute time resolution. DE 2 was on a much lower orbit, with apogee and perigee at 1012 km and 309 km.

By the time of the CDAW 9 events, some of which are discussed here, the DE 1 orbit had changed so that images were of the southern hemisphere. This study is based primarily on magnetic data from the northern hemisphere. The conjugacy of auroral forms between hemispheres may be used to remedy this situation, although such conjugacy should be used with care for both imaging and magnetic data. A case of good conjugacy, based on magnetic data, is discussed in Section 3.d. In such a case, by reversing the magnetic latitude in DE images, a northern hemisphere pseudo-image could be created. In practice, such images were of limited use in face of the greater spatial and temporal resolution from Viking images which were available for the CDAW events.

3. DMSP Imaging

The Defense Meteorological Satellite Program already referred to has consisted of a number of satellites carrying various instruments, with an imager being a constant feature due to its meteorological importance. Among the earliest satellite auroral images available, and sometimes accompanied by other data streams such as particles or magnetic perturbations, DMSP images have been used extensively in auroral studies. DMSP images are in the near infrared and visible. They are obtained in pushbroom fashion at relatively low altitude but show a nearly 3000 km wide strip of the Earth, although not all parts of the image are obtained simultaneously. During passages over the Earth's dark side the cameras are automatically adjusted to record images at low light levels, including auroral and man-made emissions. Due to low satellite altitude and the nature of the instrument, excellent spatial resolution is attained in DMSP images.

Image data for the relevant time periods being difficult to obtain, and given that for several of the events other imaging (or similar data) are used, DMSP images do not figure prominently in this study.

4. POLAR

The POLAR spacecraft carries 11 different science instruments of which the one of most interest here is the UV Imager (UVI experiment)[4]. POLAR was launched on February 24, 1996 into a 2 × 9 R_e polar orbit with a period of approximately 18 hours[5]. The CCD-based imager has a relatively narrow field of view of about 8 degrees (circular format) [Liou *et al.*, 1997]. Recent developments in UV imaging allowed the problems of contamination from longer wavelengths, which occurred with Viking, to be largely circumvented, and the camera has good spatial resolution and a nominal frame rate of one per 37 seconds. Due to the higher apogee orbit, longer sequences of images (up to about 9 hours) are available from POLAR than from Viking. From typical imaging altitudes, the resolution at ionospheric height is about 30 to 40 km. As part of the International Solar-Terrestrial Physics project, POLAR data are available in near real-time to interested parties.

---

[4] http://www.geophys.washington.edu/Space/SpaceExp/POLAR/

[5] http://www-istp.gsfc.nasa.gov/istp/polar/polar.html

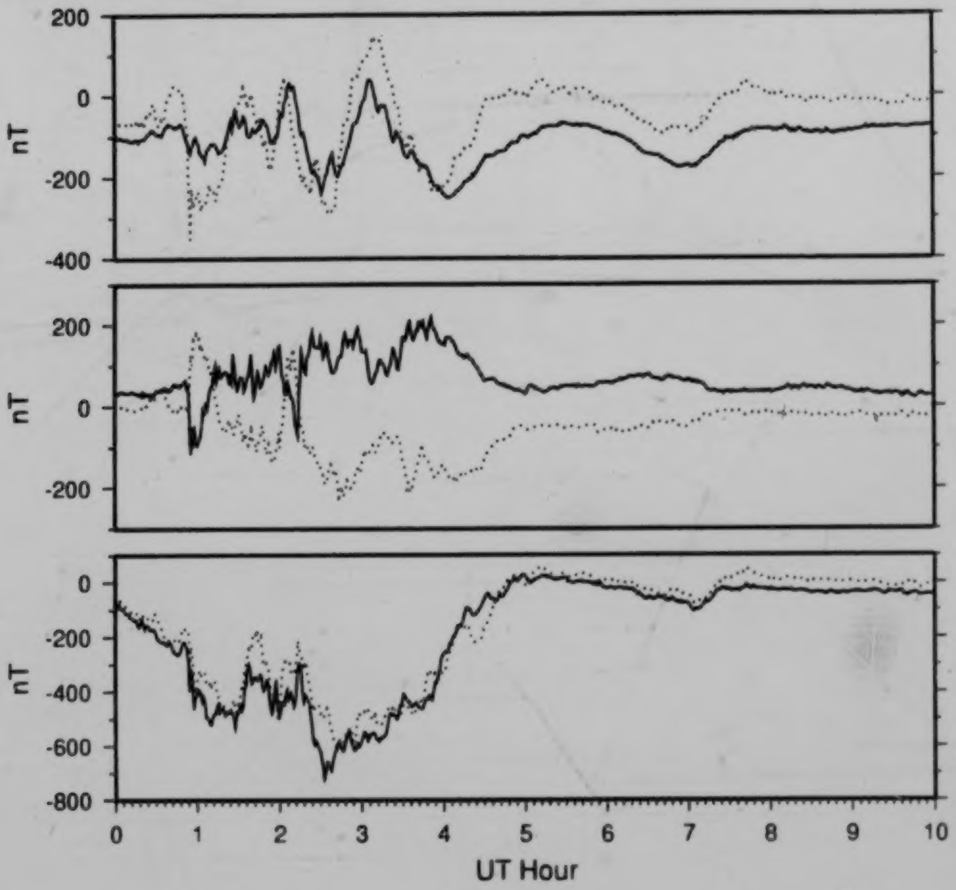
5. Meridian Scanning Photometry

Meridian scanning photometry data used are from the Meridian Scanning Photometer Array (MPA) instruments of the CANOPUS project [Xu *et al.*, 1993]. These instruments are located at Fort Smith, Pinawa, Gillam, and Rankin Inlet in western Canada (see map in Chapter 9). Meridian scans are taken through 8 filter wheels so that narrow passbands are seen centred at 557.7, 630.0, 470.9, 486.1, 480, 493.5, and 625 nm, the latter three of which correspond to background emissions, and the others to known auroral visible light emission lines [Rostoker *et al.*, 1995]. Using an assumed height of emission, the MPA data are generally presented as intensity versus latitude. In such a form they are well suited for comparison to results of the present project when used on magnetometer meridian chain data.

d. Magnetic Data and Electric Currents

The forward modelling technique used here makes a hypothetical link between observed quantities and their cause. In the case of inversion of magnetic fields, the hypothesis is that the effects observed are those due to electric currents flowing in the upper atmosphere and magnetosphere. The time scales considered are such that polarization effects (the displacement current term in Maxwell's equation) may be neglected. However, inductive effects may occur and induce currents (in the Earth most importantly) which in turn generate their own fields and complicate the relation of source currents to observed fields. Under certain assumptions, the forward model can incorporate such effects, and as used here that is done. Those assumptions may, however, be violated. This must be considered on a station-by-station basis. Earth induction is accounted for in the forward model under the assumption of an isotropic Earth, homogeneous in layers. In certain areas this may not be a very accurate representation of the true electromagnetic situation. For example, Icelandic stations have been found to have anomalous ratios of the vertical and horizontal components of the magnetic field [Hermance and Garland, 1968]. By comparison with Greenland data, taken as non-anomalous in regard to Earth induction, Icelandic data were found to have low values of Z/H in bay magnetic perturbations. Modelling indicated that the observed low ratios of Z/H could be explained by lava at 30 km depth, a not unreasonable hypothesis given the volcanic structure and activity in Iceland. For this study, this would indicate an anomalous situation not incorporated into the forward model and thus lead to inability to accurately model Icelandic data. As a control on use of Icelandic data here, it may be compared to conjugate data from the southern hemisphere. As shown in Figure 3.9, conjugacy between Husafell, Iceland, and Syowa, Antarctica, is excellent.

Figure 3.9 Conjugacy of Husafell and Syowa Stations. Husafell, Iceland data are shown by a solid line and those from Syowa, Antarctica by a dotted line. Z is at top, Y in middle, X at bottom. Time period shown is 0 UT to 10 UT April 3, 1986.

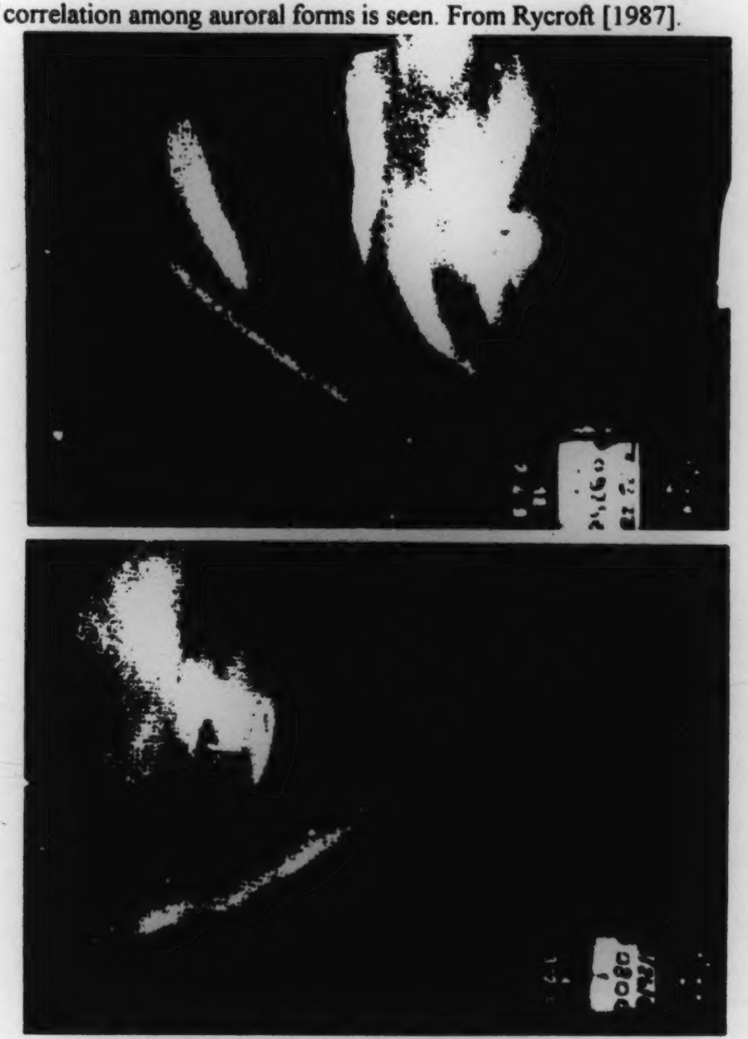


Two-second pulsation data from the XISY data set of the CDAW 9B were boxcar-averaged to one minute and referenced to a basetime of 20 UT April 2 1986 to produce this plot. The X components are almost identical. The Y components, arising from (primarily) north-south currents, are mirror images, showing that meridional current flow is oppositely directed but nonetheless highly negatively correlated. The Z components, which are sensitive largely to north-south positioning of the electrojets, are also a close match, although their sensitivity to exact conjugacy is greater than is that of the X component. Certainly no depression of Z/H ratios of a factor about 5 (as observed by Hermance and Garland) is evident. A comparison of data from Husafell and Syowa is also presented by Minatoya *et al.* [1996] shows only 'H' components, and a comparison is made to the more westerly Antarctic station Asuka. Surprisingly, although the changes in magnetic perturbations at Husafell and Syowa are highly correlated while those at Asuka appear to have a time lag, comparison of auroral forms suggested that those at Asuka were the more nearly conjugate with Husafell. The conclusion was that the conjugacy was distorted longitudinally, presumably due to field-aligned current effects. The relevance here is simply to reinforce the idea that care must be taken in use of conjugate data.

Conjugate data appear to indicate that Icelandic stations are not in fact greatly affected by deep structure. There remains the possibility of local anomalies (also often associated with volcanic regions) affecting Icelandic data. This may also be considered and dismissed for Husafell by comparing its data to those from the nearby magnetometer at Leirvogur. The magnetograms from these two stations, separated by about 20 km, were compared: the records at the two stations are basically identical and show little evidence of local anomaly. For these reasons Icelandic data will be considered as fully useable in this study.

A further use of conjugacy is to fill in gaps in station distribution. In practice there are not enough stations in the Southern Hemisphere (partly due to sparse population density and partly due to the distribution of oceans) for modelling techniques to be applied there. While the situation is better in the Northern Hemisphere, there are still regions where stations are sparse. That the high degree of conjugacy shown in Figure 3.9 is due to similarity of the physical conditions is dramatically reinforced by Figure 3.10, which shows conjugate auroral forms of a very great similarity. One might thus consider developing rules for transforming Southern Hemisphere data for use in modelling in the Northern Hemisphere. Inspection of Figure 3.9 shows that for conjugate stations simply a reversal of sign of the Y or D (east-west) perturbation is needed to make data correspond across hemispheres. This is physically reasonable as that component primarily responds to meridional currents and poleward flow (i.e. symmetric meridional flow) in each hemisphere produces a reversed sign of toroidal perturbation. The X or H component of an electrojet flowing east-west would be expected to be relatively insensitive to exact conjugacy whereas the Z component, at least near the centre of the electrojet, has a steep gradient and thus would lose coherence quickly in response to deviations from exact conjugacy. Inspection of Figure 3.1 will verify that these are indeed the observed behaviours. This aspect of conjugacy could be used in two ways: either to present more data to the modelling routine, or more conservatively to demonstrate the validity of the modelling routine by calculating fields from the model parameters and then comparing them to those observed at conjugate stations. The latter, more conservative approach is used here. Figure 3.11 shows the relatively few existing conjugate station pairs and some important regions: Halley and the north-west tip of Newfoundland; Syowa and Iceland; Siple and Roberval, Québec; and South Pole to Iqaluit on Baffin Island [Lanzerotti *et al.*, 1991].

Figure 3.10 Photographs of aurora simultaneously taken at conjugate points. A high degree of correlation among auroral forms is seen. From Rycroft [1987].



Quite apart from use of conjugate data to fill station gaps, they may be used with less weight to verify the accuracy of modelling results. Particularly at times when the existence of good conjugacy is established through use of observations such as those presented in Figure 3.1, accurate representation of the currents in one hemisphere should allow calculation of magnetic fields in that hemisphere, or, subject to the transformations and assumptions presented above, those at any postition in the opposite hemisphere. A good verification of the physical reality of parameters derived in this study would be considered to be the calculation of a conjugate station's observations from parameters derived through the modelling.

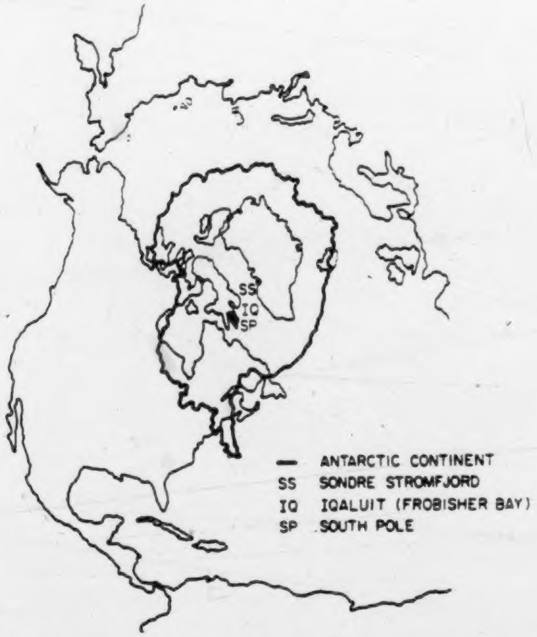


Figure 3.11 Superposed northern and southern hemisphere maps showing conjugate stations. From Lanzerotti *et al.*[1991].

It is briefly noted that the limited use of conjugacy is not restricted to this study. The AMIE procedure (see Section 4.b), which takes estimates of conductivity data as an important input, uses southern hemisphere data [Knipp, 1989, p. 23], but assigns a lower statistical weight to such cases. Conjugate data for satellites are important in this regard since conjugate data may be the only way to fill in a gap at a given time and local time. Even in cases where the satellite crosses the desired local time in the northern hemisphere, it may do so at a time when conductivities may have varied. Then the use of conjugate data for the time in question may introduce less error than use of northern hemisphere data from an inappropriate time.

e. Coordinate Systems and Transformations

A number of coordinate systems are used in geomagnetism. Certain systems associate naturally with certain problems, but while this association may simplify the problem at hand, comparison to other aspects of the problem, perhaps better studied with a different coordinate system, may become difficult.

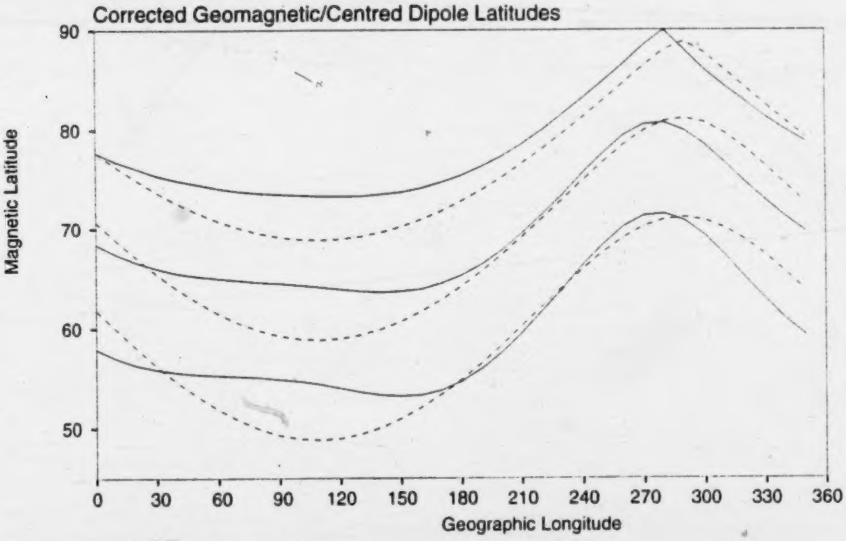
In the very good approximation that the Earth is spherical, a geodetic system related to the rotation axis of the planet and some other fixed reference point may be adopted. The standard geodetic system is based on the latitude, that is, the angle subtended at the centre of the Earth between the Equator and a point, and the longitude, which is the angle between the defined zero meridian and that of the point, as measured along the Equator. The latitude is positive in the northern hemisphere and negative in the southern hemisphere. Longitude is positive if the shortest angular distance to the meridian of a point is obtained by going eastward from the zero meridian, which is defined to be that of Greenwich, England, and negative otherwise. Longitude may also be defined as varying from 0 to 360 degrees going only eastward, in which case it is often referred to as east longitude. It is sometimes useful to speak of the colatitude of a point, which is, in degrees, 90 minus its latitude. This coordinate is analogous to the polar angle in standard spherical coordinates.

The simplest analogous magnetic coordinate system is obtained by defining polar points and a standard meridian and considering angles analogous to the latitude and longitude to be the magnetic latitude and magnetic longitude. This system is useful as the internally generated field of the Earth is dominated, in a spherical harmonic expansion in a Cartesian system based on the centre of the geodetic system, by the dipole term [Langel, 1987]. The coordinate system thus defined is referred to as the dipole coordinate system, or geomagnetic coordinate system. In order to differentiate it from other dipole-based systems, the adjectives tilted, oblique, and/or centred are often used. This will be the primary coordinate system used here and it is here referred to as the centred dipole system. A major aspect of the choice of coordinate system is that the problem should be well represented in it: operational concerns may dictate that the problem be adequately represented while not requiring undue computational baggage. The centred dipole system is orthogonal, easily related to the geodetic system by rotations, and has a simple representation of field lines along which currents may flow. Its suitability for geomagnetic modelling, at least on a regional scale, has been demonstrated in the studies of Kisabeth [1972, 1979] and Kisabeth and Rostoker [1973, 1977]. The definition of the magnetic poles required in order to define the centred dipole coordinate system in practice amounts to specifying the geodetic location of the north magnetic pole. That pole location is not constant but varies, for reasons which remain unclear (as indeed does the entire mechanism by which the Earth's internal field is generated). Although this variation has been small since about 1900, in principle the pole location should be calculated from empirical formulae, as a first step in defining the centred dipole coordinates. In 1980 the geodetic coordinates of the north magnetic pole were latitude 78.8° N, east longitude 289.3° [Langel, 1987]. By April 1986, the position as calculated from routines of Hapgood [1991] was latitude 78.85°, east longitude 289.1°. Pole positions calculated from formulas given by Hapgood [1991], appropriate to the date in question, are used throughout this thesis.

Within both of the coordinate systems defined above, vectors may be represented in local Cartesian systems. The local geodetic system is conventionally defined with an axis X directed northward, an axis Y directed eastward, and a Z axis completing a right-handed system by being directed vertically downward. Components of the magnetic induction in the local geodetic system are usually referred to simply as the X, Y, and Z components. Another common way of describing the magnetic field at a given location is to emphasize its strength and declination. In the World Data Center system, also used by the United States Geological Survey, H is the horizontal intensity in nT, D the declination (an angle), and Z the vertical intensity. Z is identical to the Z component in the local geodetic system and is positive downward. H is the modulus of the vector sum of the geodetic X and Y components. The direction of these horizontal components is expressed as the declination angle D (often given in minutes of arc or degrees) between them and the geodetic X axis, measured eastward. The conventional HDZ specification of the field does not directly specify a vector since D is in fact an angle. A useful vector quantity, due to the observed ordering of currents by the geomagnetic field, is the representation of the field in a local geomagnetic system. By analogy to the practice for the geodetic local system, the magnetic induction components can be referred to as X_M, Y_M, and Z, with the Z component identical to that of the local geodetic system. In this study, unless indicated otherwise, X_M, Y_M, and Z components are in a local Cartesian system based on the centred dipole main field model. The X_M axis is taken to be parallel to meridians in the centred dipole coordinate system, while the Y_M axis is directed magnetically eastward (lies 90° clockwise from the X_M axis when facing radially inward). Components in the local Cartesian systems are related to one another through appropriate rotations. There is a loose practice of using the terms H, D, and Z for what is here referred to as X_M, Y_M, and Z. It is important to know the units accompanying magnetic data to avoid confusion. Another variant is to use HDZ, sometimes referred to as HEZ to indicate that the eastward value is not an angle, as the axes in a truly local coordinate system established by orientation of a magnetometer on a quiet day. Such local magnetic components will sometimes be used in what follows: if this is what is supplied by a data source it is not always possible to transform to the true centred dipole (or any other) system. For modelling in such a case, it has to be assumed that the local and centred dipole systems do not dramatically differ. Alignment with local magnetic coordinates is often done for campaigns, especially for pulsation data in which absolute directional accuracy is not essential (see for example Lanzerotti *et al.* [1991]). The true declination of a site is based on local magnetic measurements, usually made through aeromagnetic surveys, and reflects higher-order terms in an expansion of the true observed field (including local anomalies due to ore bodies, etc.). Such a declination is required to convert a compass reading, for example, to true north, and is usually given on maps along with an epoch and annual variation. Global maps of declination (variation) are available (e.g. U. S. Navy Hydrographic Office [1965], or see Forbes [1987] or Langel [1987]).

The centred dipole system is based on only one term in a multipole expansion of the Earth's internal field. Field models incorporating higher orders may also be used to define coordinate systems. The most commonly used such system is the corrected geomagnetic system (CGM) of Gustafsson [1983; see also reference therein]. Like the centred dipole system, this coordinate system changes as the components of the multipole change slowly with time. Unlike that system, corrected geomagnetic coordinates do not form an orthogonal system and are in fact defined numerically. Lines of constant corrected geomagnetic latitude result from the tracing of field lines from a circle in the dipole equatorial plane to the surface. The field lines are the tangent lines of the magnetic field supplied by the chosen model field (which is usually the IGRF standard model appropriate for the epoch). While the corrected geomagnetic system has found many applications and its lines of constant latitude seem to form a good system for organizing magnetic data, here the complications involved in its use are not likely to be worthwhile. Below, certain data sets may be presented in corrected geomagnetic coordinates, but it will be made clear in appropriate places that comparisons to the modelling results, presented in centred dipole coordinates, must be undertaken with care. The two coordinate systems do not differ dramatically. Nevertheless, special care must be used in comparison of corrected geomagnetic coordinates to centred dipole coordinates in certain longitudinal sectors. This is illustrated by Figure 3.12, which shows that there may be as much as about 5° of difference between the two systems in the 60° to 120° geomagnetic longitude sector (northern Europe). While centred dipole curves are the result of a coordinate rotation and vary smoothly, between the geodetic latitude minus the dipole tilt (about 11°), to the geodetic latitude plus the dipole tilt 180° away, the corrected geomagnetic curves have a complex although basically similar dependence on longitude. The form of the curves illustrates the dependence of the corrected geomagnetic coordinates on higher-order terms in the expansion of the Earth's field: at places the corrected latitude has opposite curvature with longitude to that of the one-term centred dipole approximation.

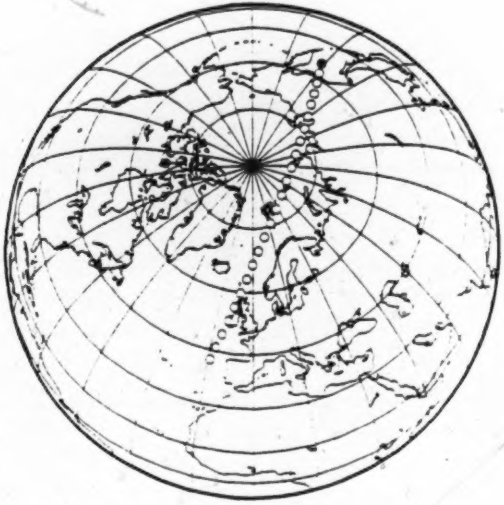
Figure 3.12 Curves of CGM and centred dipole latitudes as a function of geodetic longitude. Solid curves are CGM latitudes corresponding to geodetic latitudes of (from top) 80°, 70°, and 60°, with geodetic longitude varying. Nearest dotted curve to each is the corresponding centred dipole latitude curve.



In practice corrected geomagnetic coordinates are obtained by interpolation in a table such as that supplied by Gustafsson [1983]. Spacecraft data may be referred to the corrected geomagnetic system by tracing from the spacecraft location to the ionosphere in geodetic coordinates (since routines to calculate IGRF field values usually take as input geodetic coordinates and since spacecraft ephemerides are also calculated in that system). A transformation is then made from the ionospheric 'foot' position to corrected geomagnetic coordinates. Ephemerides of the DMSP satellite provide both geodetic coordinates and corrected geomagnetic footprint coordinates computed in this way [F. Rich, personal communication to G. Rostoker, 1991]. In this study the geodetic coordinates of the footprint were converted to centred dipole geomagnetic coordinates, which is thus a consistent system in which to present results. As discussed in section 3.a.1.c. the spacecraft data are given in a coordinate system moving with the spacecraft and must be transformed to an Earth-based system. This transformation is a rotation through an angle which varies with the spacecraft position, much as the transformation from local geodetic to centred dipole coordinates is a rotation through the centred dipole declination, which varies with location.

As more particularly discussed in section 3.a.1.c, DMSP data are supplied with an ephemeris specifying the footprint position in geographic coordinates and in the corresponding corrected geomagnetic coordinates. Since the present study uses centred dipole geomagnetic coordinates, the most appropriate coordinate system to use for comparison of spacecraft observation with results of ground-based modelling are not in fact supplied by the ephemeris. For this reason, the geographic coordinates of the footprint were transformed into the centred dipole system. It was verified that there was only a small difference between the centred dipole coordinates computed in this manner and the supplied corrected geomagnetic coordinates. Figure 3.13 illustrates the path of the spacecraft by showing the footprint positions, at one minute intervals, for a northern hemisphere pass of the DMSP F7 spacecraft on April 2 1986 (860402) from 2320 to 2350 UT. The top panel shows the spacecraft trajectory in geographic coordinates with continental outlines shown for reference. Bottom panel shows the spacecraft trajectory in magnetic coordinates with the centre of the diagram being at magnetic longitude zero, magnetic latitude 70°. Corrected geomagnetic coordinates from the ephemeris file are used to produce the path marked by white dots (initial point marked in black). The nearby grey path is the path in centred dipole coordinates as used in this study. V-shaped path at the right indicates the corrected geomagnetic midnight position during the trajectory by showing coordinate pairs of midnight longitude and spacecraft latitude. The spacecraft passes midnight about four fifths of the way through the pass as would be expected from the geographic position and the UT. For the same pass, Figure 3.14 illustrates the corrected geomagnetic latitude at one minute intervals, the corresponding centred dipole latitudes, and the smooth interpolating curve which provides centred dipole coordinates at any time. The centred dipole-corrected geomagnetic differences are within the expected range and the interpolation of coordinates works in a smooth manner and is seen to accurately reflect the ephemeris information. It is noted in passing that cyclic coordinates (longitudes or local times) must be converted to continuous functions before interpolation. For example, magnetic local time for the trajectory shown in Figure 3.13 would change discontinuously in the ephemeris from above 0 h to nearly 24 h as the spacecraft passes through local magnetic midnight. Before interpolation it is essential to convert this to a smooth function, which in this case would result in a value passing from above 0 h to negative values.

Figure 3.13 Coordinates of DMSP F7 satellite footprints on a northern hemisphere pass. Top figure shows geodetic, bottom magnetic footprints. See text for details



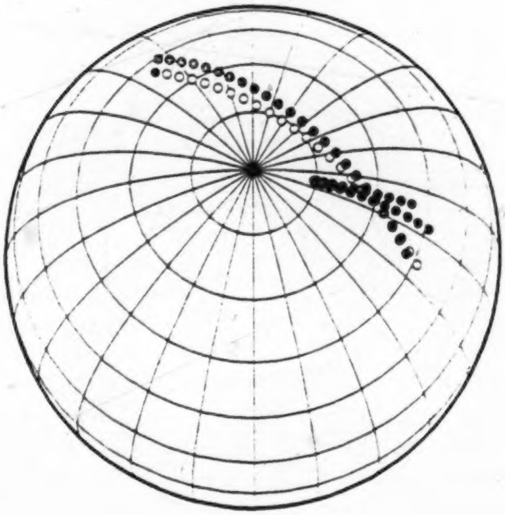
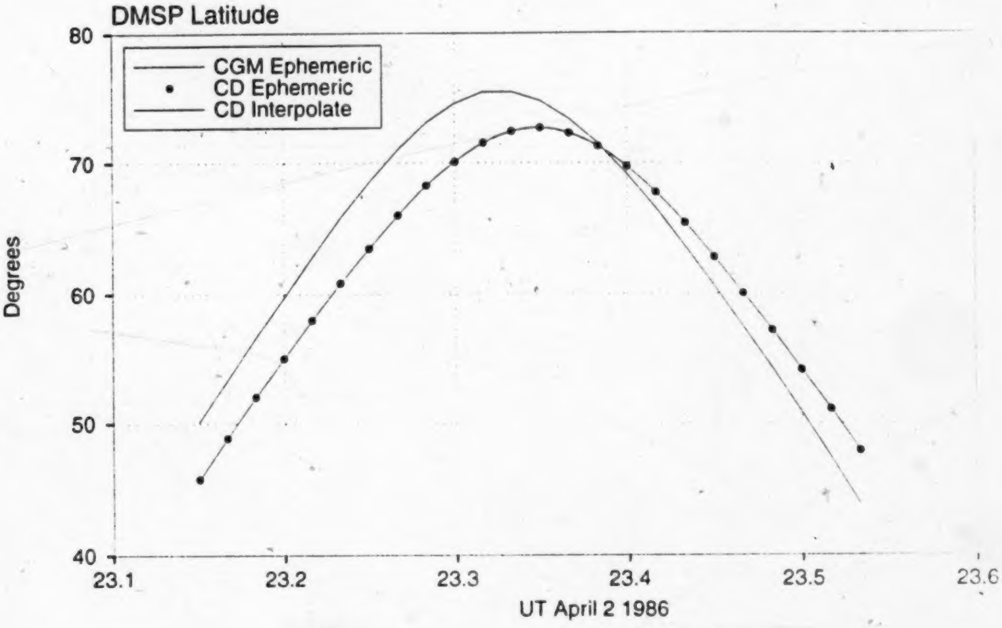


Figure 3.14 Latitude of DMSP F7 spacecraft on April 2 1986. CGM latitude of the spacecraft footprint is shown by line segments with time spacing one minute. Corresponding centred dipole latitudes, as calculated from the geodetic coordinates in the ephemeris file, are shown by dots at one minute intervals. Latitudes assigned by interpolation are shown by a solid curve which passes through the dots.



Given the role of the solar wind in forming the magnetosphere, it has been observed that magnetospheric structure and events are to a large degree ordered by the angle from the Sun-Earth line, which is conveniently expressed as a time. The subsolar point is readily calculated given the coordinates of the Sun and the orientation of the Earth. The coordinates of the Sun are obtained from a solar ephemeris while the orientation of the Earth may be expressed through the sidereal time at Greenwich. Both are standard calculations in positional astronomy [Duffett-Smith, 1990]. The antisolar point is 180° in longitude from the subsolar point and at the negative of its latitude. The geomagnetic coordinates of the antisolar point may be calculated in the desired coordinate system and the difference between the magnetic longitude of any point and that of the antisolar point, expressed in hours, is the magnetic local time (MLT). One hour of MLT is equivalent to 15° of longitude. The several types of MLT possible differ significantly only in the polar regions. In this study, unless otherwise specified, the centred dipole MLT is used.

Yet more near-Earth coordinate systems may be defined to reflect aspects of the geomagnetic field and the phenomena which it influences. Viking images are generally presented with eccentric dipole coordinates (EDFL). Invariant coordinate systems may be defined based on aspects of particle motion and are useful in meridian scanning photometry and in organizing particle data. In a dipolar field the invariant system is identical to 'centred dipole coordinates' and the two terms will sometimes be used interchangeably in what follows. In practice the advantages that the specialized coordinate systems lend to certain restricted studies will not be evident here. The many possible systems can be confusing, and here an attempt will be made to refer all data to the centred dipole system.

Brief mention must also be made of systems used for 'magnetospheric' spacecraft. The outer magnetosphere is more strongly organized by the solar wind and its interaction with the dipole than by the dipole alone. Since the solar wind flow (but not field) is largely radial, useful systems are based on the Sun-Earth line. The two most commonly used such systems are Cartesian with differences in choice of the orthogonal axes. They are the Geocentric Solar Ecliptic (GSE) and Geocentric Solar Magnetospheric (GSM) systems. Both have the X axis pointing from the Earth to the Sun. In GSE, the Y axis is in the ecliptic plane pointing toward dusk and the Z axis is pointed toward the ecliptic pole. The GSM system is useful for roughly showing the effects of the angle of the Earth's dipole in the solar wind. Its Z axis is antiparallel to the projection of the dipole into the plane perpendicular to the X axis, and the Y axis is the second component of an orthogonal triad. Transformations between these systems are described by Russell [1971], Hapgood [1991], and Appendix 2 of the recent book edited by Kivelson and Russell [1995].

Chapter 4. Modelling Techniques

The large-scale magnetic effects associated with auroral substorms may have certain of their characteristics ascertained without detailed modelling of the causative currents. Inspection of Figure 3.1, along with the knowledge that the topmost traces (down to PINA) are from stations arrayed in southward order from the north, would allow the deduction to be made that, at 0330 UT for example, westward electric currents of roughly 1 MA flowed over the array. This would correspond to a time when the current wedge was active. The pattern of midlatitude perturbations can also be used to deduce characteristics of the current wedge. A tendency of the pattern of perturbations of the magnetic D component to move westward with time was deduced by Rostoker *et al.* [1970] by examining the phase of that component across midlatitude magnetometer arrays used to study Earth induction effects. Clauer and McPherron [1974] made profiles of substorm midlatitude magnetic perturbations on a global scale and compared them qualitatively to those arising from a simple line current loop in the form of a wedge. Their analysis rested on examination of the profiles with time, and not on detailed comparison with a model. Wiens and Rostoker [1975] used symmetry arguments on plots of ground perturbation, primarily based on the D component, to study westward motion of the central meridian of the substorm current system. Fields consistent with all of these observations are those arising from the substorm current wedge much as depicted in Figure 1.20, consisting of downward field-aligned current at its eastward edge, an ionospheric path supporting westward current, and upward field-aligned current at the western edge. For qualitative studies, it is not in fact necessary to model the current system in any quantitative way to make general inferences. In this work I attempt to do such quantitative modelling for the wedge and for the driven system in order to make quantitative statements about their characteristics and relationships.

Rostoker [1975] pointed out that the wedge need not exist as a physical entity, but that a rapid westward shift of the upward field-aligned current which 'drains' the morning sector westward electrojet would produce a perturbative effect equivalent to that arising from a closed loop of wedge current. This statement was consistent with observation at the time of electron precipitation as the *de facto* signature of upward field-aligned current, coupled with the fact that a believable gradient in ionospheric conductivity could lead to upward field-aligned current where observed. On the other hand observational evidence for downward current was lacking and explanations for its origin focussed on near-Earth space which was less explored than the ionosphere. From a modelling point of view the actual current wedge region would be identical regardless of its physical nature. A quantitative analysis, as done here, may illustrate to what degree the wedge simply reflects westward motion of the upward field-aligned current associated with the electrojet and to what degree it is a separate current system.

In order to do precision modelling of the sort required to answer the above and other questions, a method of calculating magnetic fields due to current systems is required. This may, as here, be coupled with a method of optimization of parameters, or may be used in a simpler manner to justify some assertions about the behaviour of auroral zone currents.

Although preliminary work [Connors, 1993] showed that modern symbolic systems such as Mathematica® [Wolfram, 1991] could be useful in calculating the magnetic fields due to current configurations, in practice the routines earlier developed at the University of Alberta were already available in a useful form and had seen repeated use and testing. These techniques for precision solution of the Biot-Savart integral, giving the magnetic induction resulting from a specified electric current configuration, are described in detail by Kisabeth [1972]. Model results, generally from loop systems similar to the substorm current wedge, or from north-south systems, were compared to observations through the use of latitude profiles. Data from stations arrayed north-south (available, for example, from the array operated by the University of Alberta in the 1970's) were plotted as a function of latitude at a given time. Through inspection of the observed profiles, parameters of current systems could be deduced. Modelling, including effects of induced currents in the Earth, allowed changing of the parameters to obtain an optimized fit as judged by plotting the data and the modelling output on the same graph. Results indicated suitable fits to substorm perturbations in the auroral zone from wedge-like current loops where the ionospheric path followed the auroral oval [Kisabeth and Rostoker, 1977]. Similar techniques were used to deduce the importance of north-south current flows in explaining perturbations to the D component observed to be associated with propagating auroral forms [Kisabeth and Rostoker, 1973]. The wedge-like Birkeland current loops were modified by 'kinking' the middle such that the ionospheric current deviated southward before resuming a basically westward path. Here the emphasis is on simple wedge-like loops but with north-south currents (the importance of which was only beginning to be realized at the time of Kisabeth's work) included. Going beyond the use of latitude profiles near the auroral zone, with the implicit reduction of the problem to a local one in two dimensions, global three-dimensional modelling is attempted here.

a. Basic Problem

The basic aim of modelling is to use observations to produce an understanding of a physical process, through some technique linking the two. In the case of magnetic modelling the observations are of magnetic induction, the aim being to deduce the causative currents. Being able to do this uniquely is not evident. Other aspects of the understanding of the physical processes involved may be needed. In this case an important piece of physical knowledge is that electric currents arise from differential motion of charged particles and that electrons are observed to precipitate into the auroral ionosphere at times when magnetic perturbations are measured. This is evidence of the existence of electric currents which follow nearly vertical paths as opposed to paths entirely within the ionosphere. From surface magnetic measurements alone this is not able to be deduced. A long-standing controversy in space physics was that between Birkeland's and Chapman's views about the nature of the auroral current systems. Birkeland incorporated his deduction that auroral emissions were due to vertical motions of particles into his reasoning whereas Chapman's approach was more analytical but excluded, after consideration, this important adjunct information. Further aspects of this dispute, as it relates to inversion, will be discussed in Chapter 5. As undertaken here, the modelling links parameters of a simple yet physically reasonable and now accepted (Birkeland-style) 'forward model' to the observations.

Once optimized to give a best match to the data, the parameters have physical meaning which can be used to analyze the current systems. We proceed by discussing first the basis of magnetic modelling as used in the Automated Forward Modelling routine used in this work. The methods of optimization used are presented in Chapter 5. This chapter also describes (in their entirety but not in detail) other modelling techniques which are in use today.

1. Biot-Savart Integral

The Maxwell equations governing the magnetic induction **B** in magnetostatics are Gauss' law $\nabla \cdot \boldsymbol{B} = 0$ and Ampère's law $\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J}$ where **J** is the electric current density (in units of A m^{-2}). Following Slater and Frank [1947] we define the vector potential **A** by requiring $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ which constrains **B** to satisfy Gauss' law since the divergence of a curl is identically zero. Substituting this into Ampère's law and making use of a vector identity,

$$\nabla \times \nabla \times \boldsymbol{A} = \nabla(\nabla \bullet \boldsymbol{A}) - \nabla^2 \boldsymbol{A} = \mu_0 \boldsymbol{J},$$

we further constrain **A** by requiring $\nabla \cdot \boldsymbol{A} = 0$, leaving a Poisson equation for **A**. Green's solution [Jackson, 1975, Chapter 5] of this Poisson equation is

$$\boldsymbol{A} = \frac{\mu_0}{4\pi} \int \frac{\boldsymbol{J}}{\boldsymbol{r}} \mathrm{d}^3 \boldsymbol{r}$$

and taking the curl gives the induction at an observing point $\mathbf{r}_0$.

$$\boldsymbol{B}(\boldsymbol{r}_0) = \frac{\mu_0}{4\pi} \int_v \frac{\boldsymbol{J}(\boldsymbol{r}) \times (\boldsymbol{r}_0 - \boldsymbol{r})}{|\boldsymbol{r}_0 - \boldsymbol{r}|^3} \mathrm{d}^3 \boldsymbol{r},$$

where v is the region of space containing currents.

2. Earth Induction

Near-Earth magnetic field measurements reflect not only the primary field of the planet and the direct effects of currents in space, but also induced fields. These arise in the Earth since on the time scales of interest the planet is a conductor, and a time-varying magnetic induction is a source of electric field as described by Faraday's law

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$$

The electric field induced in this way is present within the conducting medium of the Earth and gives rise to electrical current. The resulting magnetic induction is detected along with that arising from the sources of interest.

The problem of the induction from a finite source above a spherical Earth was considered by Ashour [1971]. His aim was to produce relations which could be used for numerically represented fields, as opposed to those presented as a spherical harmonic expansion (which does figure in his derivation). He also gave relations for an infinitely conducting body of radius b within an Earth of radius a, and for the lower frequency case where the smaller sphere would have to be considered to be of finite conductivity. In practice his boundary condition matching method is used with a 'superconductor' located at a depth beneath the surface, that depth being chosen to be appropriate to the time scale being modelled. By variation of this depth parameter while modelling latitude profiles during active periods, Kisabeth [1972] found that an appropriate depth for the 'superconductor' during such periods was 250 km, and this value has been used in modelling in the present work. Kisabeth also rederived Ashour's results in his Appendix A2, and obtained the general form of the Biot-Savart integral near the conducting Earth. This form is incorporated into computational routines used in this study, but it is not instructive to discuss it in detail here. Rather, the general effects of uniform Earth induction are discussed and the 'anomalous' effects due to non-uniform induction in the real Earth are alluded to. The latter effects are often due to relatively near-surface conductivity variations.

In spherical coordinates with r=a defining the surface and q=b/a the ratio of the radius of the 'superconductor' to that of the Earth, the vertical component due to internal sources is

$$Z^i(a,\theta,\phi) = -q^3 Z^e(q^2 a,\theta,\phi)$$

according to Ashour, where the superscripts i and e refer to the internal and external fields, respectively. The field due to external sources, evaluated at the internal radius q^2a, will be less than it would be at the surface. Since also $q<1$ this internal contribution is smaller in magnitude than is the external vertical component. In cases of interest here, then, the internally generated field must be of opposite sign and smaller magnitude than the externally imposed field. The total field at the surface will therefore have its vertical component reduced relative to the situation in which Earth conductivity is not considered. The expression for the horizontal components is more complicated, since

$$H^i_\alpha = q^3 H^e_\alpha(q^2 a,\theta,\phi) - \frac{1}{a^2 q}\int_0^{q^2 a} r H^e_\alpha(r,\theta,\phi)\,dr,$$

where α denotes either horizontal component, for each of which the form of the expression is identical. In cases of interest here (currents near the surface of the Earth being responsible for most of the observed perturbation), the integral could be, as a very generous upper limit, one half of the previous term. The total surface horizontal components will, in the case of uniform Earth induction, be increased by an amount largely dependent on the depth of the 'superconductor'.

The effects of nonuniform crustal conductivity are yet more complex than those associated with uniform induction, and can give rise to complex anomalous behaviour. The term crustal conductivity is used in a very broad sense here to include the effects of ocean water.

which can be considerable near coasts, especially if the water becomes deep quickly. The inversion of data specifically to determine crustal conductivity is difficult in nonuniform cases and is a separate area of geomagnetic research. Here only the most general aspects of nonuniform induction will be considered, primarily to later justify attribution of lower weight in the quantitative analysis to stations which might be affected by it. The important topic of induction in technological systems also derives from the following considerations.

Since Faraday's law implies that in the conducting Earth there is an electric field if the externally applied magnetic induction (field) changes, the currents which flow due to the induced electric field may be calculated if the local conductivity is known. To first order, more current will flow where the conductivity is higher, if there are variations in the conductivity. Thus, localities underlain by good conductors should have larger current under them than nearby regions underlain by poorer conductors. In areas where layering in the Earth causes conductivity anomalies to be essentially horizontal, this larger current (horizontal) would cause a horizontal component perturbation above the conductor and a vertical component perturbation at its sides (much like the situation under an electrojet). Locally, such anomalies may be important in interpreting magnetic response.

3. Simplified Models

Since the general solution of Biot-Savart integrals, particularly with Earth induction taken into account, is difficult, simplified models abound in the literature and therefore merit some discussion. These generally rely on some approximation of the expected current system by a model system having a high degree of symmetry or infinite dimensions along one or more axes.

In the case of ground observations it is common to use some variant of 'equivalent overhead current'. In the simplest such approach, the current causing the horizontal components of the magnetic field observed on the ground is considered to form an infinitesimally thin uniform sheet at ionospheric height. The Earth and ionosphere are considered to be planar. In this approximation, the current density in A/m can be readily calculated from the magnitude of the observed magnetic perturbations and the assumed height. The orientation of the overhead current is obtained by a 90° clockwise rotation of the horizontal perturbation vector. In practice, the relation to physical units is often foregone and equivalent current vectors are used in nT, simply as a rotation of the surface field. In this way some indication is obtained of the behaviour of ionospheric currents, while the true physical currents may differ in some way [Küppers *et al.*, 1979]. An example of the use of equivalent currents in quasi-quantitative analysis of an important sequence of substorm events on March 13-14, 1989 is given by Boteler and Jansen van Beek [1992]. In computing their equivalent currents an attempt is made to compensate for induced fields (see preceding section) by first dividing observed horizontal components by 1.3, that is, the internal field augmentation of the horizontal component is taken to be 30%. The form of equivalent current used in this case is an overhead wire (so that the equivalent current really is a current and not a current density), and the external north and eastward fields (X_e and Y_e, respectively, in nT) give an equivalent current $I = 500\sqrt{X_e^2 + Y_e^2}$ at a geographic azimuth of $\theta = 90 + \tan^{-1}\left(\frac{y_e}{x_e}\right)$ degrees. Plotting locations of damage to power systems on maps, along with such equivalent currents, showed some correlation with the strongest currents at substorm onset times. The authors did to some extent use the Z component, mentioning that its sudden reversal indicated overhead passage of the electrojet. However, they felt the Z component to be unreliable due to induction effects. In certain cases, they noted, "large vertical magnetic field coincident with a zero-crossing of the horizontal disturbance field which produced anomalous equivalent currents". This combination in fact likely does not indicate a contribution from induction, but rather a strong electrojet some distance from the station. Thus, while some use of the simple equivalent overhead current approach was demonstrated, discrepancies were noted which indicate its inadequacy in understanding or even properly representing the physical currents.

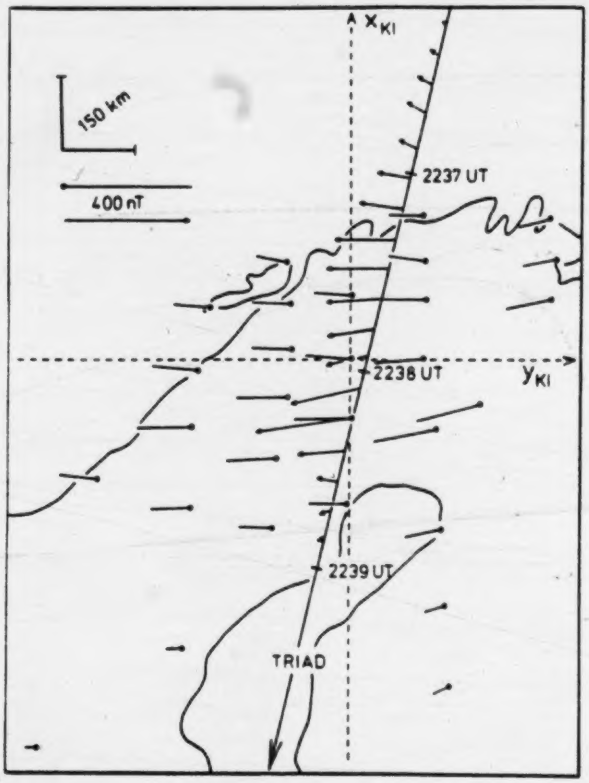
Inherent in the assumptions behind equivalent currents is that the physical current to which they roughly correspond is horizontal. Generalizing an approach based on locally horizontal currents, one may define the 'equivalent current function'. This may be defined in a purely mathematical sense, and globally as opposed to locally, as that current flowing on a spherical shell at some height (usually roughly that of the ionosphere), which would reproduce ground observations, including the vertical component. While equivalent currents dispose of the information contained in the vertical component, this is not necessarily the case with the equivalent current function. It nevertheless cannot properly represent field-aligned current flow, which has equivalent currents associated with it in the absence of physical horizontal current flow. A good example will be discussed in Chapter 7, but it might be mentioned here that, in active times, perturbations at subauroral latitudes are primarily due to field-aligned currents and that equivalent currents at those latitudes would be particularly meaningless. We defer further discussion of equivalent current functions to the section below on the KRM-AMIE family of modelling techniques, and now look at the use of simplified models for satellite data, where in fact field-aligned currents are usually the primary source of perturbations.

Boström [1964] recognized that east-west aligned sheets of current flowing parallel to magnetic field lines should be associated with the auroral oval, and proposed testing this hypothesis with a satellite experiment. By using a fluxgate detector coil set mounted on a bar magnet acting much like a compass needle in the polar orbiting satellite 1963-38C, Zmuda *et al.* [1966] were first to observe transverse magnetic perturbations above the auroral zone and to systematize their distribution relative to the auroral oval [Zmuda *et al.*, 1967]. In the first of these papers, a tentative explanation in terms of hydromagnetic waves was advanced, and in the second the authors cited Boström but did not discuss causes. Dessler [1984] claims to have been the first to suggest, as referee and editor, that transverse magnetic disturbances observed by satellite-borne magnetometers could be due to persistent field-aligned currents as had been suggested near the turn of the century by Birkeland. In any case, by 1970 an explanation in terms of Boström's hypothesis was firm, with electrons proposed as the primary carriers of field-aligned current [Armstrong and Zmuda, 1970]. Continued investigation in the early 1970s, particularly with the Triad satellite magnetometer [see Stern, 1996], led to the statistical view of the field-aligned current distribution discussed in Chapter 1 and summarized by Iijima and Potemra [1978].

These data were also successfully modelled [Armstrong *et al.*, 1975] and we return to a discussion of how this is done in the simplest (but frequently applicable) case.

It is clear that a vertical current near an essentially horizontal satellite orbital path should produce a magnetic disturbance transverse to that path. Since the regions nearest the spacecraft would contribute most to the field observed, the use of simple models based on uniform current sheets has come into wide use in interpreting satellite data. For example, the simplification of pairs of infinite and uniform Birkeland current sheets was made by Zanetti *et al.* [1983] to allow separation of the effects of Birkeland currents from those of ionospheric currents in precision magnetometer measurements from the MAGSAT satellite. The Birkeland sheets were assumed to be vertical and any radial component observed was attributed to ionospheric currents. The Fourier components of the radial component (taken as the Z component in a Cartesian system) arising from ionospheric currents are related to those of the component orthogonal to the current flow as both arise from the solution of Laplace's equation in Cartesian coordinates. Observed data was rotated about the Z axis until agreement was obtained between the orthogonal component as computed from the Fourier components and those observed. The remaining part of the horizontal perturbation was taken to be that arising from the Birkeland sheets. The effects of nonuniformities were not discussed by Zanetti *et al.* but the ionospheric currents deduced compared favorably with results from the Scandinavian Magnetometer Array (SMA) during the satellite passes. In the infinite sheet approximation the magnetic perturbation from the poloidal circuit with ionospheric currents flowing along a meridian is confined to the region between the current sheets and above the ionosphere through which the spacecraft passed. The ground perturbation from such a system would be null. In practice, even finite longitudinal extent systems of east-west aligned sheet currents, paired north-south, give little ground effect. This is largely due to the tendency of opposed field-aligned currents of the same magnitude, joined by ionospheric currents, to cancel each other's ground magnetic effects (see section 4.d.2.b). An interesting example of this is given by Sulzbacher *et al.* [1980] through the display of equivalent current vectors over the Scandinavian Magnetometer Array (SMA) with magnetic field vectors based on observations by the Triad satellite as seen in Figure 4.1. The magnetic field vectors, as observed at 800 km altitude, are essentially parallel to the equivalent current vectors. This shows dramatically that the current system detected from the ground and that detected by the satellite are not the same. Simple modelling with a poloidal (meridional) current system reproduces the satellite observations well while such a system does not contribute on the ground (nor during that part of the satellite pass which is not between the north-south pairing of east-west aligned field-aligned current sheets). Much of the SMA ground-based analysis was done with a simplified model in which Cartesian coordinates and straight magnetic field lines were used. Untiedt and Baumjohann [1993] took this approach to modelling further, and discussed the electric field in the simplified geometry. Under the assumption that **E** is irrotational, initially assumed values of conductivity and **E** led to currents whose effect compared favorably with the ground magnetic observations.

Figure 4.1 Averaged ground equivalent current and TRIAD magnetic disturbance vectors. The disturbance seen from the satellite is parallel to the currents inferred from the ground. From Sulzbacher *et al.* [1982].



Since, in satellite work, the simplifying assumption of infinite east-west sheets of Birkeland current is often made, the basic quantitative aspects are presented here. If a Cartesian coordinate system is chosen with z parallel to the main field (sometimes taken to be vertical in auroral latitudes), y eastward, and x northward, the general formula relating a perturbation $\Delta\mathbf{B}$ to the current density $\mathbf{J}$ causing it (Ampère's law $\nabla \times \Delta\boldsymbol{B} = \mu_0 \boldsymbol{J}$) can be recast as $J_z = \frac{1}{\mu_0}\frac{\partial}{\partial x}\Delta B_y$ [Potemra *et al.*, 1988]. To avoid cumbersome transformations of coordinates, study is often limited to satellite passes near a magnetic pole, so that the transverse component is very close to toroidal (examples with DMSP data may be found in Iijima *et al.* 1991 and Lopez *et al.* 1993). In later chapters of this work DMSP data are analysed without such simplifying geometry. There also, the numerical form of the current density equation will be given.

The analysis of ground perturbations in a Cartesian coordinate system proceeds in a way analogous to the use of infinite current sheets for spacecraft. It is possible to approximate the ground magnetic effects of an electrojet by a simple line current better if that line current is well above the known height of the ionosphere [Baumjohann *et al.*, 1980]. In that manner, the possible wide latitudinal extent of the electrojet may be approximated by the spreading out of the magnetic field farther from the hypothetical wire. However, while such an approach may lead to an efficient representation, it is quite unphysical and thus not desirable. More physically based modelling was performed on a combined ground magnetic-ionospheric electric data set obtained from the SMA magnetometer array and the STARE radar [Baumjohann *et al.*, 1981]. In a geomagnetically-oriented Cartesian coordinate system (with origin at Kiruna, Sweden), the ionosphere was represented by 50 by 50 km cells at 100 km height, each having height-integrated Hall and Pedersen conductivities and current. The current was related to the electric field observed (with quiet-time field removed) through the ionospheric Ohm's law (see section 2.c.2) and represented by horizontal line currents aligned with the axes. To preserve current continuity a near-vertical current (at an angle defined by the local geomagnetic inclination) was associated with each cell and its strength determined by the divergence of the horizontal currents. Both of these studies used data from a region much smaller than that active during a substorm. In the case of simplified modelling this has an advantage in that 1 dimensional modelling may suffice (as in the first example cited): this is much the same situation as the use of meridian profiles in modelling. The more physical second approach described has features in common with the KRM and AMIE methods described later in this chapter.

Magnetic perturbations observed along a magnetic meridian lend themselves to simplified modelling. A basic assumption usually made in attempting to invert such data is that currents flow along lines of constant (geomagnetic) latitude; further simplification can be had if the assumption is made that the Earth is flat and that the currents flow at a constant height [Kotikov *et al.*, 1991]. In this case the elements of the currents may be regarded as line sources and the well-known solution for the field of a finite-length wire may be applied to calculate the field. Kotikov and coworkers have used this approach to attempt to model fine structure within electrojets by solving for the currents required in many such adjacent elements to produce observed perturbations. They found that the 'resolving power' of the method was about 60 km in the ionosphere.

Another application of the line current model is its application to electrojets and as part of an automatic fitting routine [Lühr *et al.*, 1994]. Data were selected in which the eastward magnetic component was near-zero so that the current flow was likely along lines of constant latitude. The automated fitting routine attempted to optimize parameters for simple current systems consisting of a line current, a strip of constant east-west current density, or a region in which the east-west current density varied as a quadratic (parabolic model). The current densities obtained from optimizing the models using the Marquardt technique (very similar to that used in this study) were compared with those obtained by the EISCAT radar. The radar measurements yielded electron densities from which the conductivity was calculated using an atmospheric model to give collision frequencies and a magnetic field model to give gyrofrequencies (see expressions for conductivity in Chapter 2). The current was then calculated with an electric field based on F-region drift velocities. A further check was provided by rocket flights through the ionospheric current, which were available for two of the events studied. In the four days (out of four years of data) suited to analysis it was found that the parabolic model matched the EISCAT and adjacent magnetic data best. The current density estimated from the ground was found to be about 15% higher than that derived from the radar (and in two cases rocket) measurements. The discrepancy was attributed to induction effects. Further, it was found that a good solution could not be obtained by solving simultaneously for the northward (B_X) and vertical (B_Z) magnetic components simultaneously, with the B_Z zero crossing being found to be south (by 0.4° on average) of the B_X maximum. Since it was the eastward electrojet which was modelled in all cases, this displacement was attributed to the magnetic effects of a westward electrojet hypothesized to be lying north of it as would be typical in the evening sector (see Section 1.c). Since the latitude range of EISCAT is not sufficient to detect the entire two-electrojet configuration under most circumstances, Lühr and coworkers have proceeded to use their inversion technique on the longer chain which is part of the International Monitor for Auroral and Geomagnetic Effects (IMAGE) project. Results of this study are awaited.

b. KRM

The Kamide-Richmond-Matsushita (KRM) and somewhat related AMIE (see next section) techniques were developed in the late 1970's and the 1980's. Both start with conductance estimates [Knipp, 1989], and expand physical quantities in basis functions. There is provision in AMIE for varying the conductance estimates: this is not part of the earlier KRM method. Parameters in the expansion are varied until the best fit is obtained. Although these are in principle "magnetogram inversion techniques", other information is entwined with the derivation of quantities related to currents and thus to magnetic fields: this other information includes the electric fields and thus Joule heating rates. An important aspect of these approaches is that they depend to a large degree on knowledge of, or assumptions about, the conductivity. Since the conductivity is profoundly influenced by particle precipitation associated with the auroras, this dependence may be critical.

Mathematical aspects of the KRM method have been recently summarized by Kamide and Baumjohann [1993, section 4.1], and their treatment is followed here. Under the assumption that the ionosphere is a spherical conducting sheet and that steady state conditions prevail, the current continuity relation may be written as $\nabla \cdot \boldsymbol{J} = j_\parallel \sin \chi$, with **J** the ionospheric height-integrated current density, $\mathbf{j}_\parallel$ the field-aligned current density, and χ the inclination of local magnetic field lines (at ionospheric height). As detailed in Chapter 2, Ohm's law relates the current and electric field in the ionosphere and its height-integrated form is $\mathbf{J}=\Sigma\mathbf{E}$, where both **J** and **E** are two-dimensional vectors with components in the ionosphere (the field lines are assumed to be perfect conductors and thus equipotentials). In the static case $\boldsymbol{E} = -\nabla\phi$, where ϕ is the electric (electrostatic) potential. All quantities are assumed to be measured in the frame rotating with the Earth. Combining these equations results in the relation between potential and currents:

$$\nabla \cdot \boldsymbol{J} = \nabla \cdot (\Sigma \boldsymbol{E}) = -\nabla \cdot (\Sigma \nabla \phi) = j_{\parallel} \sin \chi .$$

We require ϕ to be zero at the pole and by the symmetry of the problem (at least at equinox), expect its derivative $\frac{\partial \phi}{\partial \theta}$ with respect to colatitude θ to be zero at the equator. Solution requires that a conductivity model be used as input. We must also proceed to finding the scalar ϕ by first expanding $\Sigma \nabla \phi$. While the gradient of the potential is a vector (electric field), the conductivity is a tensor and in this case of two-dimensional conductivity it is a 2 by 2 matrix, which in spherical coordinates (r, θ, λ) (we use λ to designate toroidal angle distinctly from ϕ the potential) is:

$$\begin{bmatrix} \Sigma_{\theta\theta} & \Sigma_{\theta\lambda} \\ \Sigma_{\lambda\theta} & \Sigma_{\lambda\lambda} \end{bmatrix}$$

The spherical coordinate expressions for gradient allow the components of the potential gradient to be specified as $(\nabla \phi)_{\theta} = \frac{1}{a} \frac{\partial \phi}{\partial \theta}$ and $(\nabla \phi)_{\lambda} = \frac{1}{a \sin \theta} \frac{\partial \phi}{\partial \lambda}$, so that with matrix multiplication we have

$$-\nabla \cdot \begin{bmatrix} \Sigma_{\theta\theta} \frac{1}{a} \frac{\partial \phi}{\partial \theta} + \Sigma_{\theta\lambda} \frac{1}{a \sin \theta} \frac{\partial \phi}{\partial \lambda} \\ \Sigma_{\lambda\theta} \frac{1}{a} \frac{\partial \phi}{\partial \theta} + \Sigma_{\lambda\lambda} \frac{1}{a \sin \theta} \frac{\partial \phi}{\partial \lambda} \end{bmatrix} = j \sin \chi .$$

and with the spherical coordinate expressions for divergence this may in turn be expressed as a scalar equation:

$$-\frac{1}{a^2 \sin \theta} \left[\frac{\partial}{\partial \theta} (\{ \Sigma_{\theta\theta} \frac{\partial \phi}{\partial \theta} + \Sigma_{\theta\lambda} \frac{1}{\sin \theta} \frac{\partial \phi}{\partial \lambda} \} \sin \theta) + \frac{\partial}{\partial \lambda} (\{ \Sigma_{\lambda\theta} \frac{\partial \phi}{\partial \theta} + \Sigma_{\lambda\lambda} \frac{1}{\sin \theta} \frac{\partial \phi}{\partial \lambda} \} \sin \theta) \right] = j \sin \chi$$

in which all of the conductivity components are assumed known. Further expansion and the knowledge that $\Sigma_{\theta\lambda} = -\Sigma_{\lambda\theta}$ (see Chapter 2) lead to

$$\Sigma_{\theta\theta} \frac{\partial \phi}{\partial \theta} \cos \theta + \{ \frac{\partial \Sigma_{\theta\theta}}{\partial \theta} \frac{\partial \phi}{\partial \theta} + \frac{\partial \Sigma_{\theta\lambda}}{\partial \theta} \frac{1}{\sin \theta} \frac{\partial \phi}{\partial \lambda} + \Sigma_{\theta\theta} \frac{\partial^2 \phi}{\partial \theta^2} \} \sin \theta$$
$$+ \frac{\partial \Sigma_{\lambda\theta}}{\partial \lambda} \frac{\partial \phi}{\partial \theta} + \frac{\partial \Sigma_{\lambda\lambda}}{\partial \lambda} \frac{1}{\sin \theta} \frac{\partial \phi}{\partial \lambda} + \Sigma_{\lambda\lambda} \frac{1}{\sin \theta} \frac{\partial^2 \phi}{\partial \lambda^2} = -a^2 \sin \theta j \sin \chi$$

and multiplication by $\sin \theta$ allows regrouping into a partial differential equation in the potential as a function of the angular variables:

$$A \frac{\partial^2 \phi}{\partial \theta^2} + B \frac{\partial \phi}{\partial \theta} + C \frac{\partial^2 \phi}{\partial \lambda^2} + D \frac{\partial \phi}{\partial \lambda} = F ,$$

where $A = \Sigma_{\theta\theta}\sin^2\theta$, $B = \sin\theta\left[\frac{\partial}{\partial\theta}(\Sigma_{\theta\theta}\sin\theta) - \frac{\partial\Sigma_{\lambda\theta}}{\partial\lambda}\right]$, $C = \Sigma_{\lambda\lambda}$, $D = \sin\theta\frac{\partial\Sigma_{\theta\lambda}}{\partial\theta} - \frac{\partial\Sigma_{\lambda\lambda}}{\partial\lambda}$, and $F = -a^2 j \sin^2\theta\sin\chi$. The inhomogeneous term F on the right hand side of the equation must be related to observations to permit a solution for the potential. If the ionospheric current is decomposed into a divergenceless 'toroidal' current and a 'poloidal' current derived from a potential [Vasyliunas, 1970], the toroidal field can be represented in terms of an 'equivalent current function' ψ which may be derived from surface observations [Kamide *et al.*, 1976]. The analogue of a potential representation of the toroidal current j_T is that $\boldsymbol{j_T} = \hat{\boldsymbol{r}} \times \nabla\psi$ where $\hat{\boldsymbol{r}}$ is the unit radial vector [Kamide, 1988, p. 368] and Ohm's law may be developed into $\nabla \times (\nabla\psi \times \hat{\boldsymbol{r}}) = \nabla \times (\Sigma\nabla\phi)$, where the operation of taking the curl has removed any trace of the potential term. This relation of the equivalent current function and the potential may be manipulated [Kamide, 1988, pp. 369-371] to result in an inhomogeneous differential equation resembling that found above, but with the inhomogeneous term based on observable quantities. That term now becomes

$$F = \sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{\partial^2\psi}{\partial\lambda^2}$$

which allows numerical solution for the potential. Since the conductances are taken as known in KRM, currents and other electrodynamic quantities may be calculated once the potential is known.

Once the ionospheric current is calculated in KRM, its divergence allows field-aligned currents to be computed. To study the accuracy of field-aligned currents inferred in a KRM solution based on high-latitude magnetic data, Sun *et al.* [1984] associated an upward and downward field-aligned current pair with each grid point in KRM solutions. Including equatorial closure current, the perturbations associated with these systems were computed using the algorithm for currents along dipole field lines associated with Kisabeth's method (see section d of this chapter: the Kisabeth [1979] reference was used in the calculations). It was found that D-component bays were generally well reproduced by this procedure at six low to mid-latitude (21.28° to 40.01° N eccentric dipole) northern hemisphere stations. The H component was less well reproduced, with observations often being about 15 nT negative with respect to expected values from computation. The main aim of the study was to show that the KRM method produced correct field-aligned currents, and to a degree that was demonstrated. Quite apart from the validity of the KRM procedure, the basic idea behind such a modelling project is that there should be a connection between observations at mid-latitudes and auroral-zone currents since their associated field-aligned currents cause most of the perturbations associated with low to mid-latitude magnetic bays. The present study attempts to exploit that relation in the inverse manner to that used by Sun and coworkers: rather than having a solution for auroral zone currents *a priori*, the mid-latitude perturbations are used together with those from the auroral zone to determine the parameters of the auroral zone currents.

The magnetogram inversion technique (MIT), used at the Siberian Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation (SibIZMIR), is primarily the development of V. M. Mishin. Like KRM, it has as primary inputs the ground perturbations and a conductivity model for the ionosphere, and attempts global solutions [Mishin, 1990]. In the so-called MIT-2 method, the basic KRM/MIT-1 approach is used but the conductivity may be dynamically changed during the solution. The changes are made to increase the conductivity, according to a statistical formulation, where there is upward field-aligned current in the solution. Certain derived quantities, such as polar cap boundaries, are also calculated in MIT-2.

c. AMIE

We now proceed to detail the workings of Assimilative Mapping of Ionospheric Electrodynamics (AMIE). This description of the mathematical aspects of AMIE follows Richmond [1992] and Knipp [1989] with changes to make the notation consistent and with a modified justification of the approach. The term 'assimilative mapping' which is part of the acronym indicates that the method is able, in principle, to incorporate a wide range of data types as input. Derived quantities which are related to the potential and conductivity are fit to data using a modified linear least squares fit (see Section 5.c for a discussion of standard least squares). The fit is linear in that a linear combination of expansion basis functions is used: those functions are not necessarily linear in their arguments. Individual parameters in the expansion are basically devoid of independent physical meaning. We can, for example, write that the expansion of the potential ϕ is linear in the basis functions ϕ_i with linear parameters a_i and can be expressed as $\phi(\theta,\lambda) = \sum a_i \phi_i(\theta,\lambda)$. Derived quantities which may be included in the modelling include the electric field $\boldsymbol{E} = -\nabla\phi$, the ionospheric current $\boldsymbol{I} = \Sigma\boldsymbol{E}$, the field-aligned current $j_\parallel = \nabla\cdot\boldsymbol{I}$, and magnetic perturbations $\Delta\mathbf{B}$ which may be calculated from the total current. We neglect for the moment the sources of these measurements, and instead note simply that the relationship of derived quantities to the electric potential is in all cases linear. In the original form of AMIE [Richmond and Kamide, 1988], as in the KRM method, the height-integrated conductivity Σ (which AMIE workers refer to as the conductance) is assumed known before proceeding. Possible nonlinearity in Ohm's law is removed by this assumption. Knipp [1989] explored methods of removing this limitation. For simplicity, we will first consider the case of known conductivity. Due to the linearity both of the operations involved in obtaining the derived quantities and of the expansion, expansions totally analogous to that for the potential may be written in basis functions appropriate to each observable, and having the same linear parameters a_i. Two forms of expansion can be useful: the instantaneous and the statistical. The instantaneous form corresponds to the problem of inverting a set of observations at one time, while the statistical may reflect prior knowledge, including prior sets of instantaneous observations. By this approach, deviations of the instantaneous situation from the statistical situation may be considered, which should enhance the linearity and thus the validity of a given instantaneous solution. Both the i-dimensional parameter vector **a** and the j-dimensional observation vector which we will call **b** may have their instantaneous and statistical values defined, with subscript 0 for the statistical value. If the matrix of basis functions is denoted **A**, then we can write that $\boldsymbol{b} - \boldsymbol{b}_0 = \boldsymbol{A}(\boldsymbol{a} - \boldsymbol{a}_0) + \boldsymbol{v}$, where $\boldsymbol{v}$ is a vector of errors in each measurement/expansion due to observation error and series truncation. This relation may also be viewed as an expression for the error: $\boldsymbol{v} = \boldsymbol{b} - \boldsymbol{b}_0 - \boldsymbol{A}(\boldsymbol{a} - \boldsymbol{a}_0)$, or, more to the point of what we need as a solution, as an expression to be solved for the parameters **a**: $\boldsymbol{b} - \boldsymbol{b}_0 - \boldsymbol{v} = \boldsymbol{A}(\boldsymbol{a} - \boldsymbol{a}_0)$. The solution to the latter may be written, using the formal inverse A^{-1}, as $\boldsymbol{a} - \boldsymbol{a}_0 = \boldsymbol{A}^{-1}(\boldsymbol{b} - \boldsymbol{b}_0 - \boldsymbol{v})$, which does not get us anywhere since $\boldsymbol{v}$ is not known.

It is shown in Section 5.c that for linear least squares problems in which $v = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{a}$, the solution which minimizes the norm of the error vector is $\boldsymbol{a} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{b}\boldsymbol{A}^T$, with the superscript T denoting a matrix transpose. In Chapter 5, possible weighting of errors at stations is incorporated into the procedure in scalar form, but Knipp [1989] gives a solution for the case where these statistical properties are known beforehand as an error covariance matrix $\mathbf{C}_v$, with elements $\boldsymbol{C}_{vij} = \langle v_i v_j \rangle$, the statistical expected value of the error products. These errors are reasonably assumed to be uncorrelated, so that the expected value of products is zero except in the autocorrelative case where i=j. $\mathbf{C}_v$ is thus a diagonal matrix with value v_i^2 on the diagonal in row i. The inverse $\mathbf{C}_v^{-1}$ is simply a diagonal matrix with the inverse of these values in corresponding places. The weighted least squares result, which is identical to that shown in Chapter 5 but which shows the error matrix explicitly, is $\boldsymbol{a} = (\boldsymbol{A}^T\boldsymbol{C}_v^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{C}_v^{-1}\boldsymbol{b}$. This solution derives from minimization of the quantity $(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{a})^T\boldsymbol{C}_v^{-1}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{a})$, called the *cost function*. This is an equivalent problem to the minimization of weighted chi-squared discussed in Chapter 5, but with the error covariance matrix plainly shown. To constrain an AMIE solution to be near past solutions an extra term is added to the cost function, which is essentially a weighted deviation of the proposed model parameters from those which are statistically found. This term is simply $(\boldsymbol{a} - \boldsymbol{a}_0)^T\boldsymbol{C}_s^{-1}(\boldsymbol{a} - \boldsymbol{a}_0)$ in an analogous fashion to the original cost function term, with $\mathbf{C}_s$ the statistical expected value of the products of error $(\boldsymbol{a} - \boldsymbol{a}_0)$ elements. This is easily written as $\boldsymbol{C}_{sij} = \langle (\mathrm{a}_i - \mathrm{a}_{0i})(\mathrm{a}_j - \mathrm{a}_{0j}) \rangle$. In principle, one should obtain both error covariance matrices from the analysis of statistical properties of large amounts of data and of the deviations of many solutions from the 'real' solutions, but this is not possible. Noting that this problem is discussed in detail by Richmond[1] and Kamide [1988], we bypass it and now simply state that a constrained minimization which incorporates prior knowledge is needed for the modified cost function (with statistical averages shown explicitly in both terms)

$$(\boldsymbol{b} - \boldsymbol{b}_0 - \boldsymbol{A}(\boldsymbol{a} - \boldsymbol{a}_0))^T\boldsymbol{C}_v^{-1}(\boldsymbol{b} - \boldsymbol{b}_0 - \boldsymbol{A}(\boldsymbol{a} - \boldsymbol{a}_0)) + (\boldsymbol{a} - \boldsymbol{a}_0)^T\boldsymbol{C}_s^{-1}(\boldsymbol{a} - \boldsymbol{a}_0)$$

That minimization has solution $\boldsymbol{a} = \boldsymbol{a}_0 + (\boldsymbol{C}_s^{-1} + \boldsymbol{A}^T\boldsymbol{C}_v^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{C}_v^{-1}\boldsymbol{b}$, which is the central equation of AMIE. Before returning to discussion of the covariance matrices, brief mention will be made here of the incorporation of conductances into the model as a variable [Knipp, 1989; Richmond *et al.*, 1990]. Conductances can be determined by ground-based radar studies and such observed conductances, if available, may be used to augment statistical knowledge by modifying the input conductance model at a given time. Information about conductance may be obtained from past knowledge of particle precipitation and its effect on the ionosphere. The particle precipitation measurements from DMSP series satellites are useful in this respect, especially since this is ubiquitous instrumentation (unlike the magnetometers, which are not present on all of the satellites). This appears to be the favored source of direct data recently [B. Emery, D. Knipp, personal communications, 1995]. Established statistical relationships between ground magnetic perturbations and radar-observed conductances allow use of the observed magnetic fields to alter the (input) conductance pattern used. Such 'assimilation' of conductance information is done at a stage prior to the actual solution of the AMIE equation.

---

[1] An earlier but closely related discussion is provided by Richmond and Baumjohann [1983].

The KRM method has as its basic magnetic input equivalent currents. An equivalent current at a single station is generally envisaged as the current density of an infinite sheet which, at ionospheric height over the station, would produce the horizontal (external) perturbation observed. It should be immediately noted that this definition of equivalent current ignores the vertical component of the field and thus loses information about the physical system. Nevertheless, there are situations where this approximation may correspond to the physical system: examples would be the field below the centre of a relatively wide electrojet, or the field caused by the S_q physical current systems (as discussed in section 3.a). In practice, equivalent currents are often schematically depicted simply by rotating the observed horizontal perturbation 90° in the counterclockwise direction. In a plot of such equivalent currents, regions of expected high true current density generally become apparent: a disadvantage is that regions where the perturbation is not primarily due to overhead current may appear to have overhead current. The AMIE procedure, on the other hand, uses equivalent current functions rather than equivalent currents. Equivalent current functions, as discussed above, can include the vertical perturbation component. It is claimed to be the first global modelling procedure to use equivalent current functions with instantaneous data [Knipp, 1989, pp. 86-87]. Induction effects can be pronounced in the vertical component (see Section 2.c.3) and attempts are made to compensate within AMIE. The vertical component is thirty-minute averaged to eliminate high-frequency response to changes in the field. In regions suspected of coastal anomaly, low weight is given to the stations. It will be seen in work presented later in this study that induction effects appear, generally speaking, to be adequately taken into account by the Kisabeth [1972] routines which incorporate uniform induction as discussed in Chapter 2. It is likely that useful information in the Z component is totally ignored by the KRM method, and that AMIE could profitably use less time averaging. A more subtle point about use of the Z component in these routines follows.

In methods which are based on spherical harmonic decomposition, the behaviour of the magnetic components may require quite different degrees in the expansion. Apart from a normalizing factor and the dependence on order for the variation with longitude, such expansions invoke a dependence on the associated Legendre functions $P_n^m(\theta)$, where m is the order, n the degree, and θ the colatitude [Kamide *et al.* 1976]. The form $P_n^m(\theta)$

referred to here actually is a shorthand for a functional dependence which is $P_n^m(\cos\theta)$ (see Jackson [1975], p. 85, or compare equations 6.1 and 4.1a in Kamide [1988]). Commonly this shorthand is extended by denoting cos θ by the variable x. The associated Legendre functions are derivable from the Legendre functions (which are polynomials) since [Mathews and Walker, 1970, Chapter 7]

$$P_n^m(x) = (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x)$$

and (Rodrigues' formula)

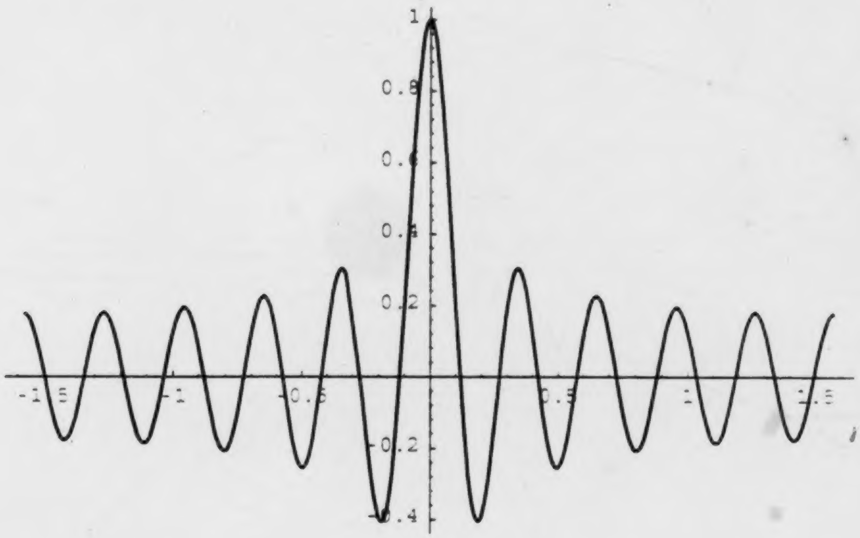
$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

It is readily seen that the associated Legendre function $P_n^m(\cos\theta)$ is, at most, of polynomial order 1+n-m. The polynomial order is highest when m=0, in which case the associated Legendre function is identical to the Legendre polynomial. The associated Legendre functions and the Legendre polynomials are complete and orthogonal in the interval $-1 \leq x \leq 1$ and although from what precedes it does not rigourously follow, the roots of these functions lie in this interval also. Thus in the highest order case $P_n^m(\theta)$ has n+1 roots in $-1 \leq \cos\theta \leq 1$, that is to say in the full range of colatitudes $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. The most general expansion of a function f of the angular variables is written, in analogy to a Fourier expansion,

$$f(\theta,\phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_{mn} P_n^{|m|}(\cos\theta) e^{im\phi}.$$

where the A_{mn} are normalization factors. The range of colatitude from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ is covered by such an expansion but there can be only at most n+1 roots in that range for any given term of the expansion. In a similar way to the limiting Nyquist frequency of a Fourier expansion, this limited number of roots limits the highest frequency angular variations which can be represented by the expansion. It is not as easy to define the equivalent of the Nyquist frequency since the roots will not be equally spaced (which they are in the case of the trigonometric functions used in a Fourier expansion). Nevertheless, it is clear that such an upper limit will exist if n, the expansion order, is finite. Since the Z component from long east-west electric current systems has a maximum and minimum on either side (north and south) of the current system, whereas the H component has only one extremum due to such a system, it can be inferred that proper representation of Z would require an order roughly twice as large as does proper representation of H. Figure 4.2 illustrates a Legendre polynomial to clarify the points made here. Although the periodicity of roots is not constant, ranging from 0.15195 between the first two positive roots to 0.15320 between the largest ones, it is very nearly so, to the point that the analogy with the Fourier Nyquist frequency may be clearly envisaged. If this function were the highest order basis function in an expansion, each positive peak could be thought of being the smallest feature possible in the expansion. Since there are ten such peaks in the entire range of colatitudes, which is π or 180°, the smallest representable feature has width roughly 18°. In practice the KRM method [Richmond and Kamide, 1988] has order 22, so one would expect widths of about 15° in latitude to be approximately the limit of the technique. This is wider than the canonical 5° average width of the electrojet and associated H perturbation. Z variations would normally have a characteristic scale of roughly half this, or about 3°. Clearly with an order limited to 22, the Z variations cannot be adequately represented. Richmond [1992] claims that in fact, with 121 expansion terms, the electrodynamical quantities can be represented down to 'scales of about 2.5° in magnetic latitude and 15° in magnetic longitude', with effects due to smaller-scale fields (in principle represented by truncated terms) entering as errors. Part of the reason for the smaller scale size achieved in practice is the use of basis functions complete over a smaller colatitude range than the 180° assumed above for the Legendre polynomials. AMIE uses a basis set complete to 56° [Richmond, 1992] and thus having a colatitude range of 34° rather than 180°, a factor of roughly five increase in potential angular resolution. This is consistent with the conclusions above and with the stated latitude resolution.

Figure 4.2 Legendre polynomial $P_{20}(x)$ for $x=\cos\theta$ in the interval $\theta = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$



Beyond the limitations imposed by finite order, an expansion based on spherical harmonics shares the property of nonlocality with Fourier techniques. Returning to Figure 4.2, the reason that this presents a problem may be seen heuristically. Imagine a data set consisting only of a single pulse at about θ=1, and near-zero elsewhere. Since the basis function shown indeed has a peak at this point that pulse, if of about the right width (we suppose that it is), will be well represented by the contribution from this basis function

However, all of the other peaks in the basis function are also introduced into the expansion approximation. Since the data set is near-zero away from the pulse, clearly these other peaks are not desired. However, the other terms of the truncated expansion contain only lower frequencies and can never cancel them out completely. In this sense, a spherical harmonic expansion cannot well represent a local feature, or conversely it can do so only at the cost of not representing the overall function. To avoid this problem, the KRM method uses as basis functions generalized (non-integer order) Legendre functions [Haines, 1985] transitioning to smoothly varying functions at a specified colatitude. Generalized associated Legendre functions are also used in magnetic field expansions attempting to represent the field only within the 'polar cap' of specified angular radius by spherical cap harmonic analysis (SCHA) [Walker *et al.*, 1995]. Since only the high-latitude regions (in practice the transition colatitude is 34° in KRM or polar cap radius 40° in SCHA) show significant discrete features, these may be represented without forcing the required high spatial frequencies (which in any case are not high enough in practice to represent the electrojet) to also be present at lower latitudes. Since a realistic profile of perturbations would often resemble the single-pulse function discussed above (with the 'pulse' being the electrojet) with few discrete features in the polar cap, one could wonder if instead three regions are not needed in such an attempt to patch up the expansion, and if the division into a high-latitude and low-latitude zone might not force oscillations to spuriously occur in the highest-latitude parts of the former.

The recent regional application of a combination of KRM and AMIE described conceptually by Sato *et al.* [1995] possibly alleviates problems alluded to in the previous two paragraphs. In this approach, the (equivalent) current function in a localized region is represented by an expansion. The local gridding is 1° in latitude by 2° in longitude rather than the 1.7° by 10° used by AMIE. The potential is solved using boundary conditions supplied by a global AMIE run and conductances from DMSP and EISCAT radars, and the other electrodynamic parameters are then derived from it. Although there is general agreement with KRM global results, the local method attains higher latitudinal resolution, resolving structures in the potential over about a 5° scale size rather than roughly 10° in the corresponding KRM result. The electric field also shows structures of latitudinal extent a few degrees, which are verified from DMSP (the F8 spacecraft) measurements. Since the technique was applied in the morning sector and the conductances were assumed to not vary with longitude, results are essentially one-dimensional, showing variation primarily with latitude. The most likely reason for the increase in resolution is not the grid size but a presumed more adequate representation of higher spatial frequencies due to the localized nature of the expansions. It is interesting to note that the mathematical techniques of AMIE were used in pre-AMIE studies by Richmond and Baumjohann [1983] to assist in representation of the field near the Scandinavian Magnetometer Array but without inversion. In these studies a Fourier series in Cartesian coordinates was used. Inversion using such an approach should be possible but may not have advantages over Sato's technique now that computational efficiency considerations are not of importance compared to human costs of implementation of a new method. Boundary conditions would still need to be supplied much as in Sato's approach.

Both the KRM and AMIE methods attempt to use magnetic data in the context of broader knowledge as to what constitutes a 'reasonable' solution. The basic approach is put succinctly by Knipp [1989, p. 9], while discussing conductivity information in AMIE: 'The constraints, in the form of *a priori* information essentially stretch the information content of sparse observations and cause the estimated patterns to adhere to statistically reasonable values in regions of limited data'. This is obviously a completely reasonable approach given that we can do little about the existence of limited data, at least in the historical approach. However, it is wise to bear in mind that limitations of these methods (AMIE and more particularly KRM) may be related to the imperfection of the *a priori* information. Particularly this is the case for the conductivity, which is an important physical variable with a very wide range of values, which are affected by precisely the phenomena under investigation.

A further difficulty with both KRM and AMIE is that the field-aligned currents are derived from the divergence of horizontal currents. This has been mentioned above, as has the verification by Sun *et al.* [1984] that the KRM field-aligned currents are basically correct. Knipp [1989, p. 89] points out that since the basic quantity is the electric potential ϕ, the field-aligned currents are actually calculated from this as $-\nabla \cdot (\Sigma \cdot \nabla \phi)$, so that both the conductance Σ *and* its gradient are important to determining them. This further highlights the great importance of good conductance estimates as input to these routines.

d. Kisabeth's Modelling Approach

An obvious method of solving a problem is to find the answer. Kisabeth [1972] presented methods of calculating the perturbations due to certain types of current systems, and proposed combinations of those current systems which reproduced the input data, which were most often in the form of perturbations observed along meridian lines of magnetometers. The proposed current systems constitute a 'forward model' and to the extent which it agrees with the data, some confidence can be placed in the physical parameters of the model. In this forward modelling approach determination of parameters was left to the modeller and several methods of resolving the Biot-Savart integral could be employed depending on circumstances [Kisabeth, 1979]. In practice, application of forward modelling to complex situations involving arbitrary distributions of stations was difficult. In this sense, an element of 'trial-and-error' [Kamide and Baumjohann, 1993, p. 98] was involved, and in complex situations the number of parameters involved, and their response at the many stations needed to uniquely optimize them, became problematic. The thrust of this study is to take forward modelling techniques, most particularly those based on direct precision integration of the Biot-Savart integral, and apply an automated technique for finding an optimal set of parameters to model magnetic data. The parameters are physical parameters and their variation is taken to indicate physical changes responsible for the observed perturbations. In practice the parameters used are simple geometric ones in this sense the method is complementary to other methods which necessitate, to a certain degree, assumptions about physical parameters under study (the conductivity, for example).

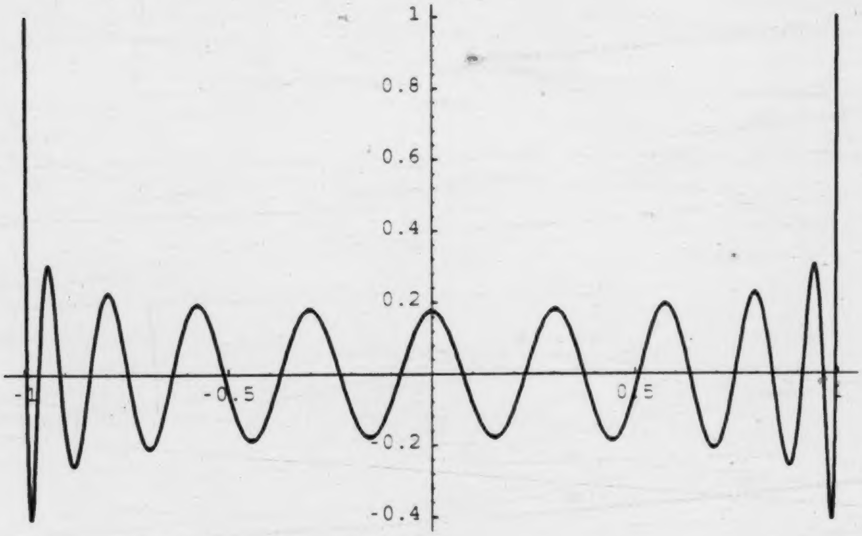
1. Gaussian Integration

Although the Biot-Savart integral for the magnetic induction may be solved analytically for certain simple geometries, it is not, in general, feasible to solve it for the currents which form a realistic three-dimensional forward model[2]. The integral must be solved numerically, and Gaussian integration [Abramowitz and Stegun, 1965, p 887] was found by Kisabeth [1972] to be an efficient method for its evaluation. This approach for the Biot-Savart integral has drawbacks, mainly due to its representation of current distributed along an integration path by point centres of magnetic induction sources. The form chosen by Kisabeth is strictly speaking the (most commonly used) Gauss-Legendre form [Press *et al.*, 1992, Section 4.5]. The point centres are positioned along the integration path where the Legendre polynomials $P_n(x)$ have zeroes, where x is the integration variable after the range of integration has been scaled to the interval (-1,1). Figure 4.2 illustrated the zeroes of $P_n(\cos\theta)$, and as a function of the angular variable θ these zeroes were seen to be very regularly spaced. In Gauss-Legendre integration, however, the basic variable is not angular, but rather the integration variable directly. The result is that the point sources of current are not evenly distributed along the path of integration. At a sufficient distance from the sources a good approximation to the exact solution is obtained. Closer to the sources which approximate the real distribution, there enter errors of approximation. This problem is typical of most numerical methods of integration, but here it bears discussion due to the fact that 'sufficient distance' may be difficult to attain for global current systems such as are used here. More particularly, the irregular spacing means that the needed distance from the path of integration varies with position. Figure 4.3 assists in understanding this problem. It shows the same twentieth-order Legendre polynomial as does Figure 4.2, but as a function of the argument directly. It is seen that the roots are more closely spaced in the peripheral regions than in the centre. As a result, the field approximation arising from numerical integration using point sources located at the positions of roots would be more reliable near the ends than near the centre. In practice, if the distance between pairs of roots as scaled to the current path becomes a significant fraction of the distance from them to the observation point, errors will become significant. This problem is particularly acute in Automated Forward Modelling, which necessitates approximate calculation of derivatives of the field with respect to parameters as part of the optimization procedure (see Chapter 5). Kisabeth [1979] considered other techniques which could be used when the 'sufficient distance' criterion could not be met and, in particular, those applicable to satellite observations, which may be made extremely close to the current system[3]. Connors [1993] used techniques based on linear current elements and analytic derivatives which preclude the errors referred to above. Here further reference will not be made to those techniques: instead Gaussian integration alone will be used with due regard to error terms. Use of short current systems (as in Kisabeth [1972]), or effective splitting of longer systems as done here, can circumvent the problems. In general, these problems do not arise for field-aligned currents, which are nearly vertical near the auroral oval, since the close spacing near the end of the path of integration occurs near the ground observation point.

---

[2] An attempt to model mid-latitude perturbations during substorms formed a preliminary study [Connors, 1993] to this work. The approximation of a flat Earth and straight segments carrying current, both in the ionosphere and 'field-aligned' was used. In that approximation, analytical solutions of the Biot-Savart integral could be used, and their derivatives calculated (using *Mathematica*®). Within the limitations imposed by the model, reasonable solutions for the parameters were obtained, and good reproduction of input data could be obtained using them.

Figure 4.3 Legendre polynomial $P_{20}(x)$ for x in the interval [-1,1].



2. Components of Forward Model Current Systems

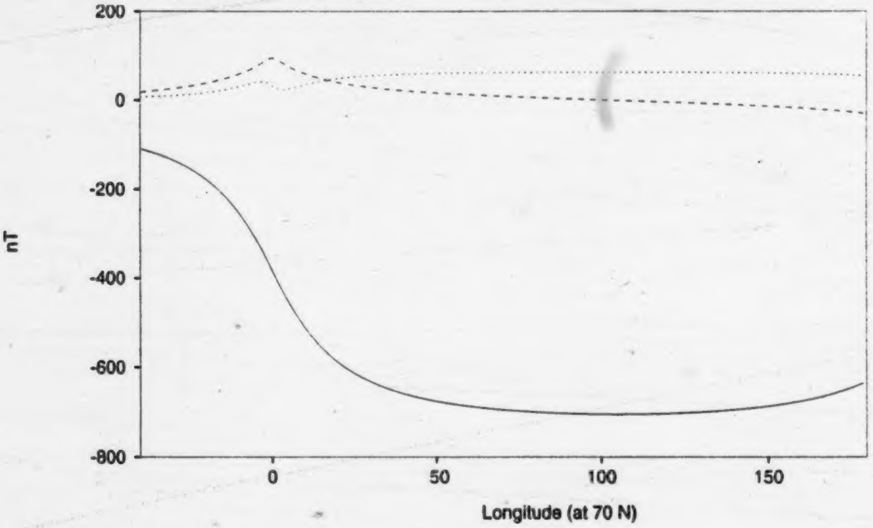
It is the aim of this section to present illustrative forward models and their associated magnetic perturbations. These are the building blocks from which the Automated Forward Modelling technique grows, when combined with a method of determining those values for forward model parameters which allow optimal representation of a set of data.

It is useful to divide the auroral current systems into two categories which produce distinct ground effect and which correspond basically to the two types of ionospheric horizontal conductivity. In doing so the terms 'toroidal' and 'poloidal' will be used in their strict sense meaning aligned parallel to latitude lines and parallel to meridians, respectively. The toroidal currents extend in longitude over several time zones but without very much north-south component of the current. Throughout much of the length of such a system the ground effect is essentially that of an east-west current overhead, and thus is mostly in the north-south component. Much of the current flow in such systems is due to the Hall effect and they are often referred to as Hall currents. The poloidal currents flow parallel to magnetic meridians and over relatively short north-south distances. As a result the effects of the field-aligned currents which feed them cannot be neglected: these effects essentially cancel those of the north-south overhead currents. As previously discussed (sections 1.d, 2.c.1, and above in this chapter), these currents were proven to exist only through satellite observations. Despite the fact that the total current in the poloidal system may exceed by a large factor that in the toroidal system, the ground effect is small and ambiguous. By analogy with the popular introductory example in which the magnetic induction from an infinite solenoid is nill, the field of the combined field-aligned and ionospheric currents in this case is said to be 'solenoidal', although this analogy is not exact. The ionospheric flow in such systems is associated primarily with the Pedersen conductivity and they may be loosely referred to as Pedersen currents.

---

[3] It should be noted that Gaussian integration can be used even in such a case. A transformation to angular variables allows this but at relatively large computational cost.

Figure 4.4 Magnetic perturbation at 70° N due to 1 MA of westward current flow in a 180° system. Magnetic components are, in this and following figures, indicated by a solid line for X_m, a dashed line for Y_m, and a dotted line for the Z component.



a. Toroidal Currents

The toroidal currents are often observed to span many hours of local time with relatively limited north-south extent and with electric current flows which are essentially parallel to lines of constant geomagnetic latitude. As discussed in the first chapter, westward current flow is commonly seen in the morning and late evening sectors, and eastward current flow in the evening sector.

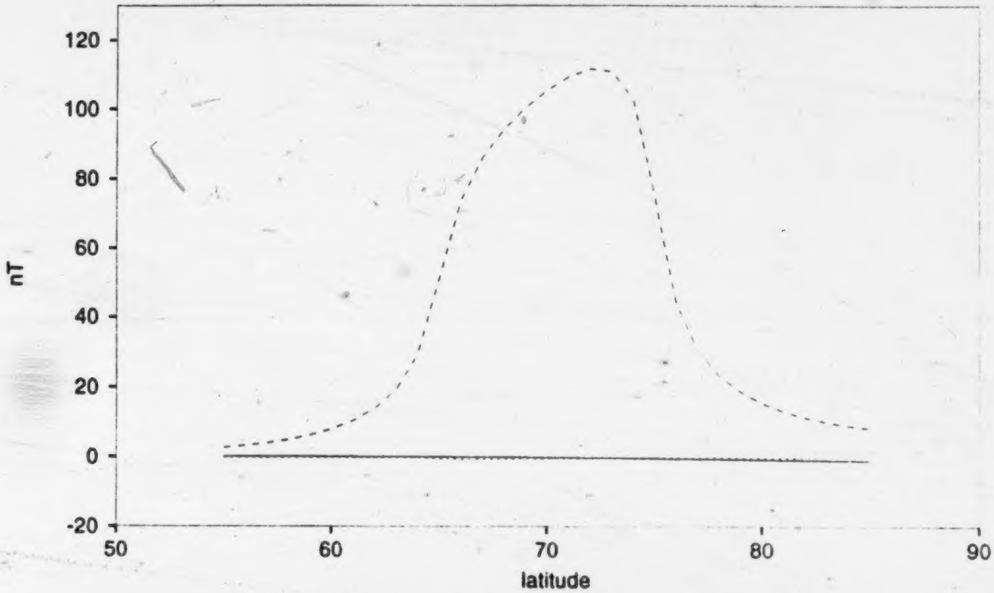
Figure 4.4 presents the perturbations associated with 1 MA of current fed into the ionosphere at 180° longitude flowing westward along an ionospheric path with 10° of north-south extent centred on 70° N (i.e. spanning latitudes 65° to 75° N), and then leaving the ionosphere as field-aligned current at 0° longitude. It will be noted that almost all of the perturbation effect is in the X (north-south) component, which has a value slightly less than -600 nT for much of the path length. The other components are much smaller in this longitude profile at the latitude of the centre of the current, and are clearly associated with the field aligned currents. The ground effect of this type of current system is a clear indication of overhead toroidal electric current flow. Since auroral zone electric fields are normally north-south, east-west current flow is generally associated with the Hall conductivity (section 2.c.2). This leads to the commonly-heard claim that 'Hall currents are seen on the ground'. More properly, the statement would be that 'only the poloidal magnetic field...generated by the electrojet proper, is observable on the ground' [Untiedt and Baumjohann, 1993].

b. Poloidal Currents

Contrary to the case with toroidal currents of large east-west extent, poloidal currents cannot in general be very long in the direction of current flow, due to the limited north-south extent of the auroral oval in which conductivity is high enough to permit significant ionospheric currents to flow. As a result, the field-aligned current effect largely cancels ionospheric ground effect in poloidal systems of large east-west extent similar to those mapped out by Iijima and Potemra ([1978]; see also section 1.d). Poloidal systems of short east-west length can produce large perturbations on the ground, as for example in the case of Ps 6 magnetic 'pulsations' (see Connors and Rostoker [1993] and references therein). However these are small-scale current systems and the focus here is on global systems (although the methods presented could be equally well used to model small-scale systems in a region where ground data on the appropriate scale is available for the period of interest). Ironically, we must consider cases where the largest systems are very difficult to detect while smaller ones may not be, and in fact may be a source of 'noise'. We proceed to consider why the perturbation from a poloidal system of large east-west extent cannot be readily detected from the ground.

Figure 4.5 shows the perturbation associated with a poloidal system in which 1 MA of current enters the ionosphere at 75° N and exits at 65° N after flowing due south through the ionosphere, this for the longitude range 180° to 0° (compare morning sector current systems in Figure 1.11). The results are shown as a latitude profile at 100° longitude, near the central longitude of the current system. The perturbation is eastward (+Y$_m$) as would be expected for southward current flow overhead. However, the magnitude is less than that which results from the same 1 MA flowing east to west as discussed above. This is due to the spreading out so that 1 MA covers 180° rather than 10° in the direction transverse to the current flow.

Figure 4.5 Latitude profile for ground effects associated only with ionospheric current in a north-south system. The current enters the ionosphere at 75° N and exits at 65° N and totals 1 MA as in subsequent figures. Profile is at longitude 100°.



The next two figures (4.6 and 4.7) show the 'jumps', again primarily in the Y$_m$ component, associated with passing under the northern and southern feed field-aligned currents. Figure 4.8 shows their sum: the total perturbation due to field-aligned currents, which is seen to be essentially the mirror image of that due to the ionospheric current, and the following figure (4.9) shows the results of this sum, almost total cancellation of all perturbations. It is thus certainly true that a poloidal system of long east-west extent produces small perturbations compared to those of an east-west system of similar extent, such as those shown in Figure 4.4. There is nevertheless a perturbation present, and one which has a trend to become larger toward the pole. It appears that, perhaps non-intuitively, the ground effect of a north-south system is mainly an offset in the toroidal (Y$_m$) field component, and that this offset is not closely localized near the current. A global view will aid in examining this effect.

Figure 4.6 Jump in Y_m component associated with a long 1 MA downward sheet current at 75° N. It is characteristic of unbalanced current flow into the ionosphere

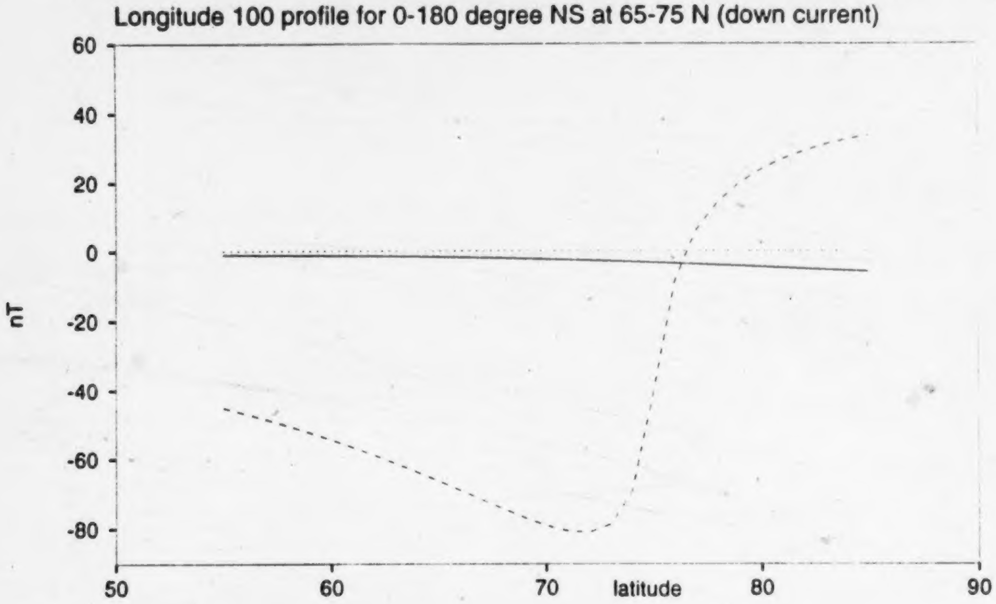


Figure 4.7 Jump in Y_m component associated with a long 1 MA upward sheet current at 65° N. It is characteristic of unbalanced current flow out of the ionosphere

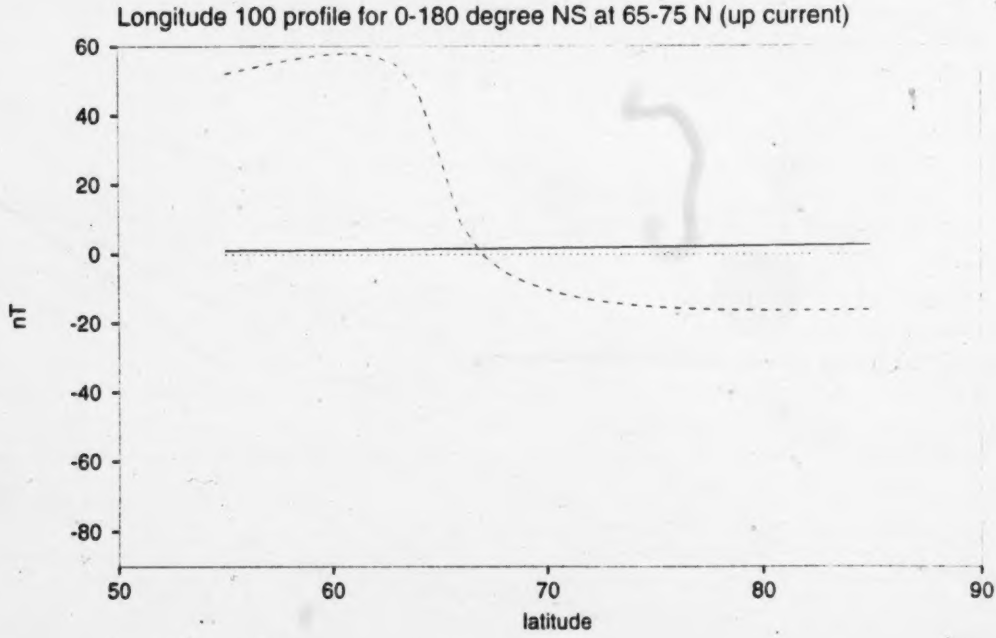


Figure 4.8 Total perturbation associated with field-aligned sheet currents of previous figures. Compare to Figure 4.5: perturbations are nearly opposite.

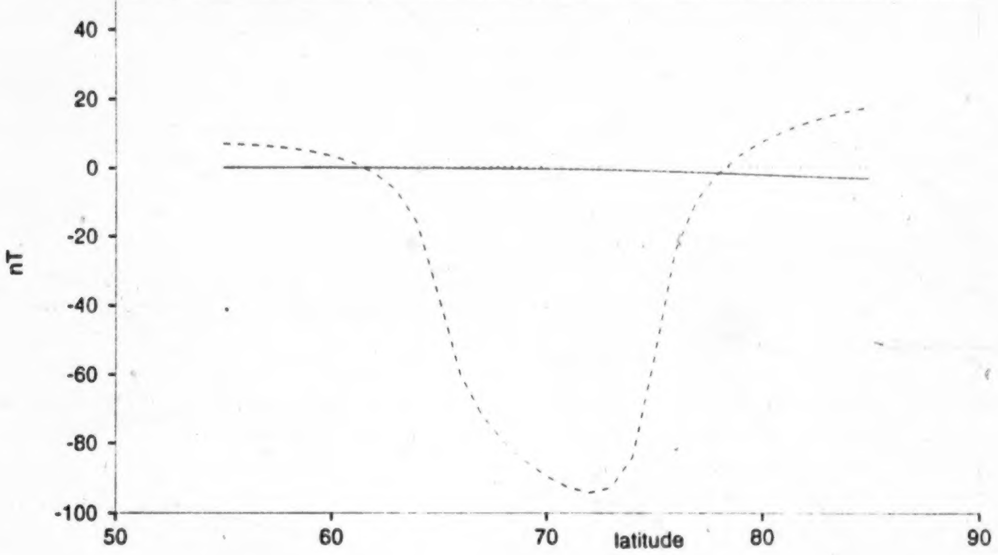


Figure 4.9 Latitude profile of total perturbation associated with north-south current flow of 1 MA with ionospheric path from 75° N to 65° N.

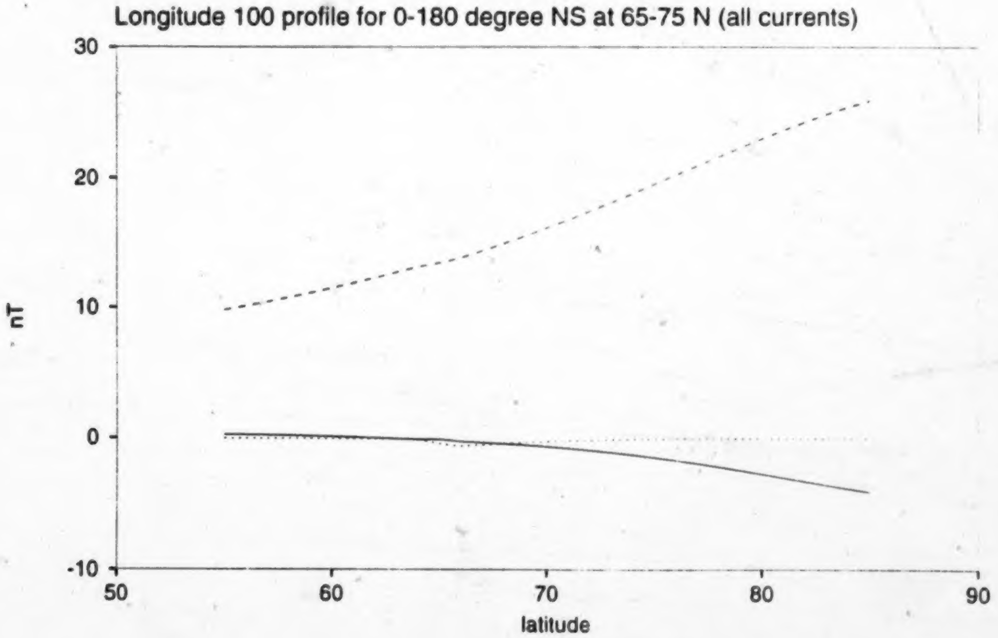


Figure 4.10 shows the effects of two 160° systems with current flow as in the Region 1/2 currents and illustrates that these systems should be detectable on the ground through enhanced polar cap fields in (bearing in mind the actual orientation of Region 1/2 currents symmetrically around solar noon) the sunward direction. The next figure (4.11) is a comparable one from Kisabeth [1979] illustrating this same result but with more current and calculation based on a different (magnetic charge) technique. For illustrative purposes, Kisabeth used current densities of 1.0 A/m at 67.5° latitude, which is larger than that usually inferred by a factor of about 10. These results both indicate that the Region 1/2 currents should give a sunward magnetic field in the polar cap. The polar cap magnetic perturbations are of interest partly because barium release experiments by Heppner *et al.* [1971] showed that the electric field there could not be related to them through any reasonable conductivity tensor. They suggested that what was seen was the effect of net field-aligned currents, in or downward in the 0800-1200 local time range and out or upward in the 2000-2400 local time range. Chen and Rostoker [1974] noted that D magnetic perturbations were not systematically observed near the proposed location of the inferred net field-aligned currents. They suggested instead that polar cap magnetic fields might be due to the distant effect of field-aligned currents threading the auroral oval. The results presented here, which are consistent with observations of the large-scale Region 1/2 currents [Iijima and Potemra, 1977], tend to support that hypothesis. Chen and Rostoker indicated the need for 'considerably more detailed modelling' to explain polar cap perturbations in terms of three-dimensional currents. The results here indicate that field-aligned currents may be included and their parameters determined (perhaps with low accuracy: see Chapter 5) if sufficient polar cap data are available.

Figure 4.10 Ground effects of two 160° long north-south current systems with current directions of Region 1/2 currents. Ground effect maximizes in the polar cap

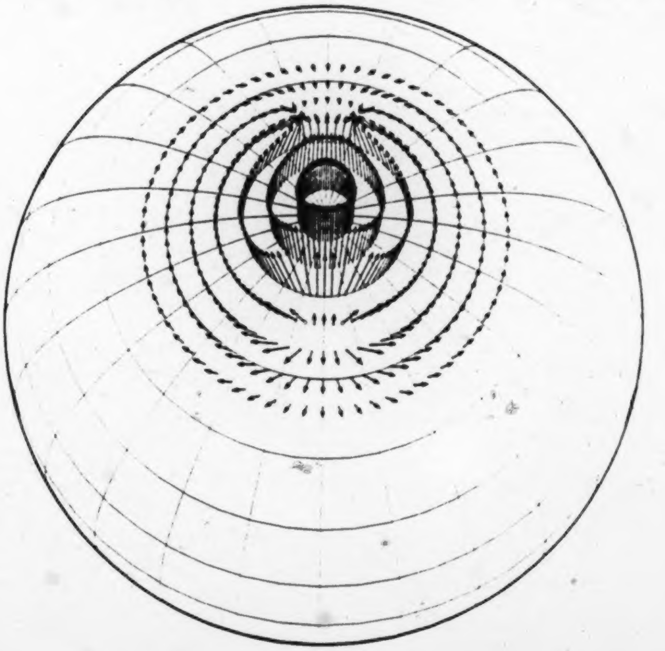


Figure 4.11 Ground effects of two 160° long north-south current systems with current directions of Region 1/2 currents as calculated by Kisabeth [1979]. Scale indicates perturbations of 200 nT.

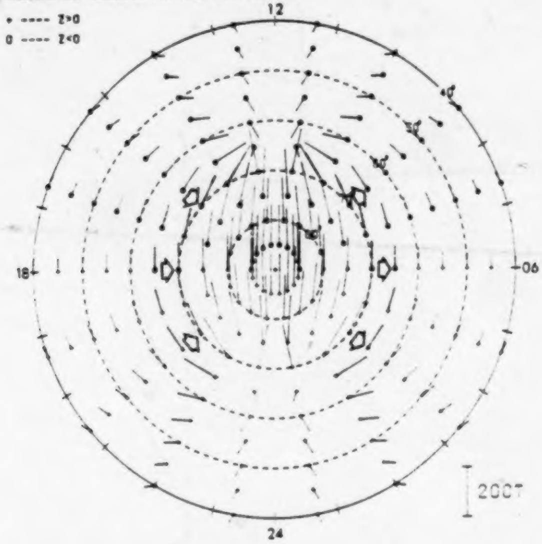
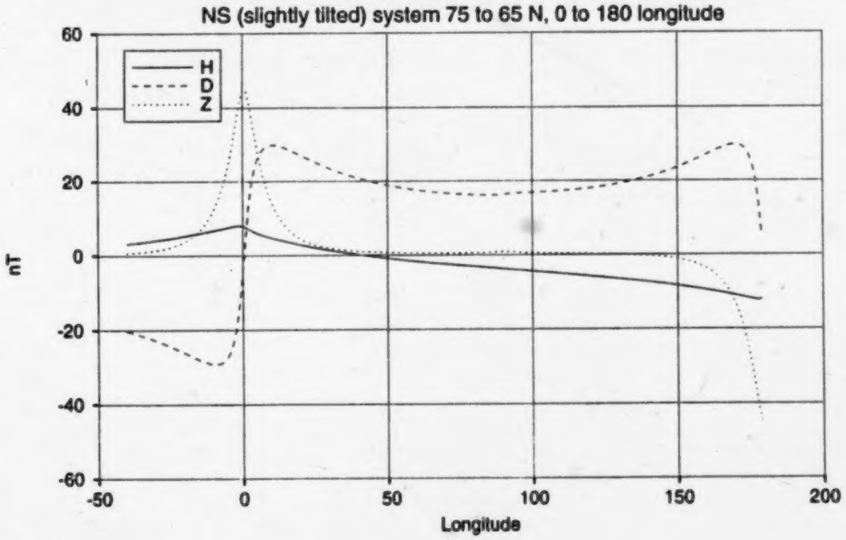


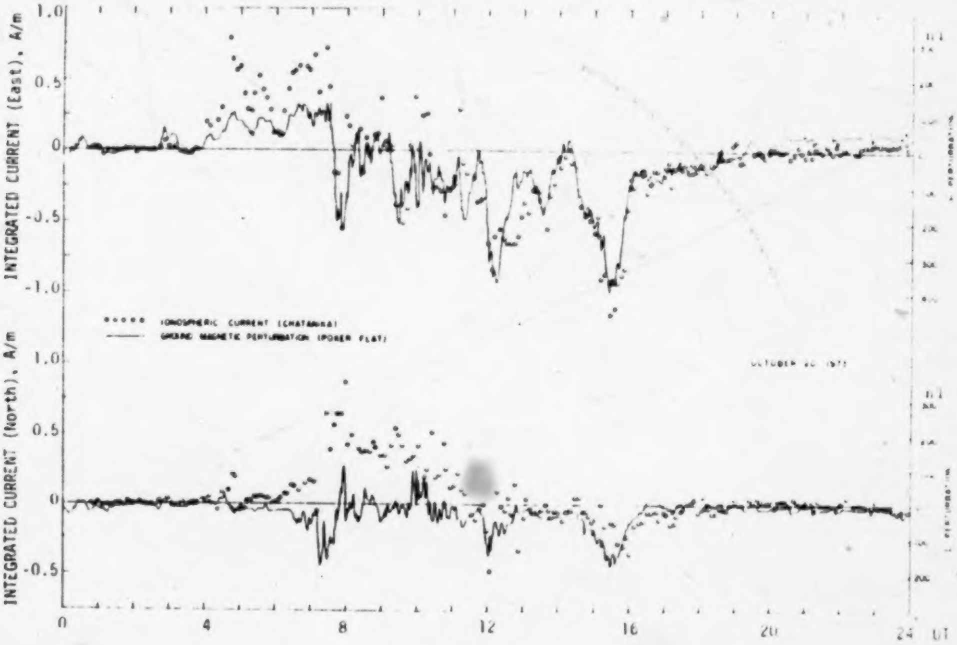
Figure 4.12 Ground perturbations associated with a long north-south current. Main effect is a small Y_m perturbation which changes sign at the end of the current system.



Ironically, directly below poloidal currents may be the poorest place to infer information about their parameters through magnetic observations. The best place to expect to do so would be at the ends of the current system, as illustrated by Figure 4.12, in which it is seen that there may be a 'leakage' Z effect and a sign reversal of the Y_m (marked as D) component.

Figure 4.13 provides further evidence of the difficulty in detecting poloidal current flow from the ground. In this figure, current density along north-south and east-west axes as determined from radar at Chatanika, Alaska, and the ground perturbation at nearby Poker Flats[4] along these axes, are compared. A generally excellent agreement is found for the eastward current and the H (northward) perturbation, while that of the northward current with the D (eastward) magnetic perturbation is markedly less good. In particular, the interval 8 to 11 UT has consistent and large northward current (as would be expected in this local time sector from the ionospheric path linking normal evening Region 1/2 currents) while little D perturbation is seen on the ground.

Figure 4.13 Ionospheric current inferred from radar compared to ground perturbation. From Kamide and Brekke [1975].



---

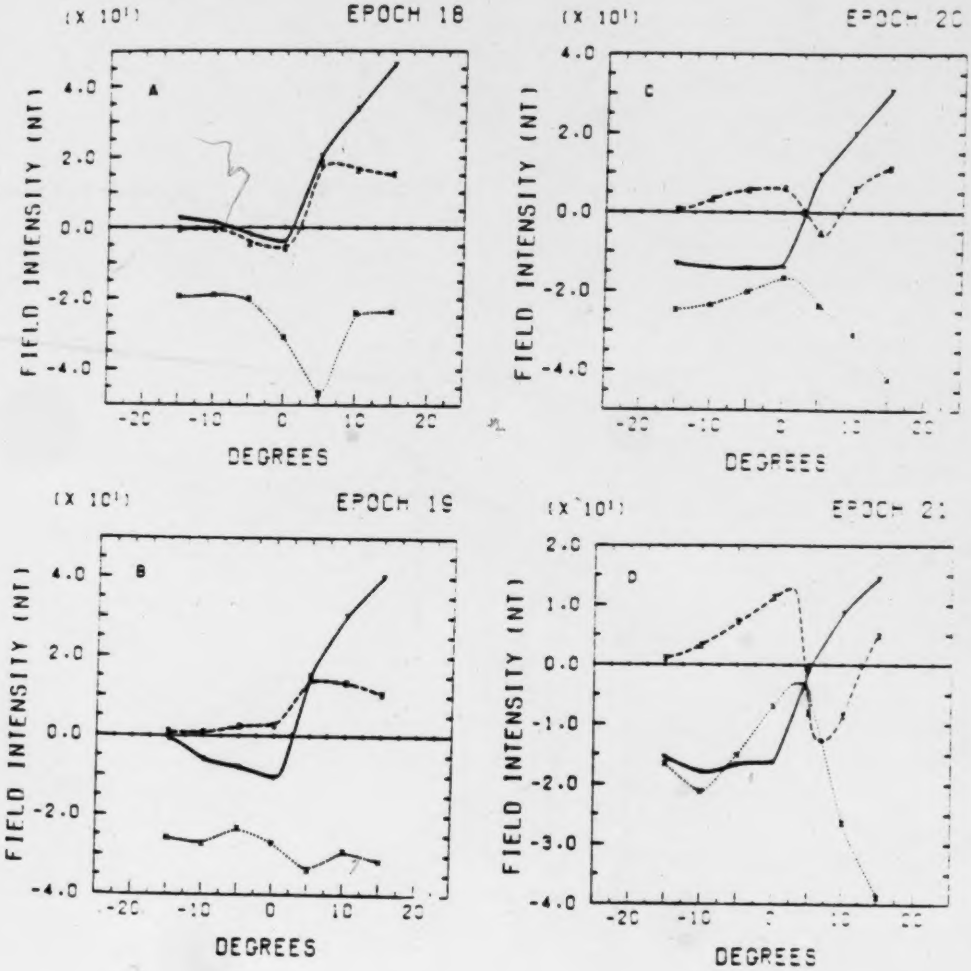
[4] 3.8 km northeast of the radar site

3. Application of Manual Forward Modelling

Sophisticated magnetic forward modelling to date has made use of the ability to calculate magnetic perturbations based on model parameters. These may be initially estimated by inspection of magnetograms, and then modified manually to optimize the fit of the model to data. This method has been notably successful for basically one-dimensional situations such as latitude profiles (Kisabeth [1972], Hughes [1978]), and less so when complex situations in three dimensions have been studied (Connors *et al.* [1991]). Geophysical inverse theory has also been applied to latitude profile data. This is a tractable problem since its application with data varying along one dimension is through one integral in that dimension (see Hughes [1978], Appendix II, or Press *et al.* [1992], Chapter 18). A numerical inversion is generally possible in such a case. The use of latitude profiles has often been very productive since many aspects of auroral current systems vary primarily with north-south position over several hours of local time. Much information can be obtained from observatories providing simultaneous observations along meridians. This fact is reflected in the global distribution of observatories, with several north-south 'chains'.

Latitude profiles can provide information about the global distribution of auroral currents if the assumption can be made that these were steady in time. In that case the rotation of the Earth carries chains eastward, providing latitude profiles at many local times. For example, Hughes [1978] used data from 10 stations along the 300° E magnetic meridian for a statistical study of the winter auroral zone. Results of superposed epoch analysis of about three months of data are shown in Figure 4.14. To compensate for variations in the actual latitude of the auroral oval and statistically reinforce common features, latitudinal translation of the profiles was done according to a distinctive feature in the profiles. In the case of the noon sector data shown, this feature was the 'base' of the Y_m component increase toward the north. These data may be compared to Figure 4.6 and it may be readily seen that a reasonable explanation for the observed Y_m component behaviour is a net downward field aligned current in this sector. The other components do not, however, all show the behaviour of the computed profile. Those farthest from noon are readily understood. The 10-11 MLT (epoch 18) profile shows a peaked negative X_m perturbation with a Z which changes sign. These are consistent with the perturbation from an westward ionospheric current. Those at 13-14 MLT (epoch 21) show the opposite behaviour (although in both cases the X_m component baseline is negative), consistent with an eastward ionospheric current overhead. The overall interpretation is one of net downward current in the noon sector and more careful inspection would show that this is generally to the north of the ionospheric currents which carry it away toward the west in the morning sector and toward the east in the afternoon-evening sector.

Figure 4.14 Superposed epoch results of Hughes [1978] for the noon sector Y_m component trends upward to the north (right) in all cases and is marked by a solid line X_m component is shown dotted, Z dashed. Ordinate value markings are in 10 nT units



Chapter 5. Optimization of Models

The parameters of model auroral current systems were described in the previous chapter To model the currents at any given time, these parameters must be adjusted to provide an optimal match to the available observations. Once that optimization is done, it is claimed that the data have been inverted within the framework of the proposed forward model As will be seen in later chapters, the simple forward models proposed do, in fact, allow a good representation of the auroral current systems. Their parameters are also demonstrably amenable to optimization based on ground based data, at least in the case of east-west systems. After a brief discussion of aspects of the forward models which are relevant to the optimization, optimization techniques themselves are discussed and demonstrated below

Inversion of geophysical data will sometimes not be able to give unique results. If a particular model is put forward to interpret certain data, there may exist other models able to fit the data about equally well. The classic relevant example of this is the long debate (1930's to 1960's) about the nature of the current systems producing geomagnetic disturbances, which has already been alluded to. As pointed out by Akasofu [1991], Chapman's assertion, based on the knowledge at the time, was that only equivalent current systems (in the ionosphere) could be discussed. Despite the economy of Birkeland's assertions at about the turn of the century that three-dimensional systems would explain not only the magnetic variations but the precipitation of particles needed to create the aurora. Chapman's view of ionospheric currents being the physical currents became accepted. This was not solely due to Chapman's dominance in the field and refusal to openly debate Birkeland's assertions, and the supportive modelling of particle trajectories done by Störmer, were considered by Chapman, and even illustrated in the influential book *Geomagnetism* [Chapman and Bartels, 1940]. The contemporary, and incorrect, context was the view, prevalent at the time (and still held by semi-informed laymen today), that particle streams ejected by the Sun were directly involved in production of aurora Particularly over the long distances this implied, it appeared to Chapman that a stream of only one charge could not hold its integrity, and that precipitating particles, although present, were in neutral streams [Chapman and Bartels, 1940, Chapter 24]. Such neutral streams would of course have little or no magnetic effect. Despite vigorous championing of the Birkeland hypothesis by Alfvén, it was only in the 1960's that the existence and importance of extra-ionospheric currents was widely realized [Dessler, 1984]. As already described, wide acceptance of the three-dimensional nature of the current systems occurred only after space measurements of magnetic fields which could be plausibly due only to field-aligned (or nearly so) currents. The basic model used here is derived from knowledge since ascertained about the physical current systems. The same techniques could, in fact, be used to give a representation of the data based on a non-physical model (such as Chapman equivalent currents). Thus, the applicability of the model here, and the uniqueness of the solution, rest on knowledge which is not incorporated into the technique. In this sense the numerical techniques are hopefully being used in keeping with Hamming's [1986] maxim that "The Purpose of Computing is Insight, Not Numbers". Nevertheless, in practice the results of the computation are numbers, and these are in general the numerical parameters which are varied in the proposed model until that model represents the data optimally

a. Parameters

The choice of parameters for a model may be viewed as part of the 'direct problem' for a good model the parameters will allow the reproduction of the data [Sabatier, 1993]. In fact the basic problem faced by the analyst is the 'inverse problem' in which data (usually contaminated by noise) are available and from which the parameters are to be deduced. Generally we will be convinced that the number of parameters being used is far smaller than the actual number required for a complete description of the physical system. For example, in the lower magnetosphere, the current density in SI (MKS) units is a sum over the charge-velocity product of at least thousands of particles. The resulting magnetic field seen at some point is physically due to the complex motion of those thousands of particles, yet may be well represented in practice from knowledge of just an integrated quantity, the current density. The very large-dimensional space of physical states and defined parameters contains such integrated parameters as a subspace. The projection of the state vector of the overall system onto the subspace defined by the parameters is the best approximation possible, within that subspace, to the overall state. The approach to approximation as viewed from a vector space point of view is discussed briefly by Anton [1987, ch. 7]. The fitting of data by models is one step removed from the fitting of functions by other functions since in fact the actual state is not generally projected onto the parameter subspace, but rather the state maps into a data space and the model provides a method of mapping the parameters into that space also. Within that space there exists an optimal choice of parameters bringing the two projections as close as possible. The choice of a model with its associated parameters and the techniques for obtaining the best values of those parameters are related but distinct. Due reflection must be given to the choice of a model. Not only the physical basis of the model may be relevant but also its efficiency. Does the data set provide sufficient information to distinguish the number of parameters chosen?

The usefulness of parameters and their choice depends on their intended application. A spherical harmonic expansion such as the IGRF representation of the Earth's main field [Langel, 1987] has important practical applications, yet most of the parameters do not have an independent physical meaning. Representation of the active fields at high latitudes by similar expansions [Walker *et al.*, 1995] can also not claim to provide parameters having a direct correspondence to physical parameters, although the spatial distribution of approximations to some physical fields (for example equivalent ionospheric currents) may be represented in terms of the parameters derived to represent the magnetic field. The approach of representing fields in terms of expansions does not yield an immediate *mechanistic model* [Bates and Watts, 1988, p. 67] for the phenomena being studied. Such a model, in its most powerful form, has both fewer parameters in general than does a non-mechanistic model, and more meaning attached to those parameters. A useful illustration of the difference may be given by considering the Earth's nearly dipolar field. In practice at least 50 parameters are used in the IGRF, of which the first is the dipolar term. If the field actually were that of a dipole, the statement of this fact in terms of the IGRF as presently used would have all parameters but one equal to zero. This may seem a trivial example, but it may be taken one step further. Consider a shift of the dipole away from the origin of the expansion (in fact the Earth's field is largely that of such an 'eccentric dipole'). In this case the field would require all terms of the finite number in the IGRF expansion for its description and even then that description would be only an approximation. Yet in fact the difference between this field and that of a centred dipole can be represented by at most only three more parameters representing the origin translation. Obviously the expansion is wasteful of parameters, those parameters having (individually) little relation to the physical situation, and provides little insight into that situation although possibly describing it well. In terms once more of vector spaces, there may well exist a vector space in which relatively few parameters suffice to describe a physical situation. In the case of particles and currents previously described, that efficient space may be the one comprised of the variables usually associated with MHD; in the case of the eccentric dipole, it is clearly that of the dipole parameter and translations. In this study the aim is characterization of substorm current variations in a manner which leads to physical insight. For that reason a parametrization is sought which adequately describes the data (reflecting success of the optimization process) and also has parameters with physical meaning. The aim here, then, is to invert data into a small and efficient subspace in which the parameters have physical meaning and shed light on causality. This point is reflected in the recent statement about the substorm current wedge (SCW) current system, from Tsyganenko [1997]: "due to the very dynamical nature of the disturbances and a relatively small number of available simultaneous data, one can expect to fit meaningfully only a few parameters of the SCW, reflecting its most important characteristics. This also requires the model to be relatively simple."

A further aspect of having a small number of parameters, each having a direct physical interpretation, is the practical aspect. If the physical effects of changing the parameters are understood by the analyst, the analysis itself can be more effectively done. To quote Bates and Watts [1988, p. 72], "one of the best things one can do to ensure a successful nonlinear analysis is to obtain good starting values for the parameters - values from which convergence is quickly obtained." Parameters which lend themselves to initial analysis and choice of starting values by graphical methods, by understanding the response of the goodness of fit to their derivatives, or by insight into behaviour with reduced dimensionality (i.e. holding one or more parameters fixed), are preferred.

An attempt at simple characterization of substorm current systems was made by Cramoysan and Orr [1993] within the framework of the McPherron current wedge. The parameters chosen are only four, the meridians of filamentary field-aligned currents, the latitude of the ionospheric path joining them, and the magnitude of the current. The geometrical parameters are considered to be fixed while the current varies. A slightly more refined, but still limited, approach to this problem allows all of the parameters to vary, yet only considers one wedge to be active. An independent but similar approach was used by Connors [1993] and considerably earlier by Horning *et al.* [1974]. The Cramoysan and Orr approach arguably oversimplifies the problem by having not only one simple current system in the form of the substorm current wedge, but further by having only one parameter (the current) changing in association with it. That the other parameters of the wedge generally also vary is well known (see e.g. Wiens and Rostoker [1975]). The use of more parameters in the other cited studies allows a good fit to the data, but in both cases only stations distant from the auroral zone current systems were used, so that information was lost and the parameters solved for correspond to 'lower-order' approximations.

In general, and in particular if data from the auroral zone are included, more parameters must be used, as there is more to auroral zone currents during a substorm than simply the substorm current wedge. In practice the global current system during active times is regarded as having two other loops active which may be reasonably represented by a similar overall configuration to that of the wedge. These loops are those of the eastward and westward electrojets in the evening and morning sectors, respectively. They are similar to the substorm current wedge in that evidence exists of regions of net downward field-aligned current feeding them, ionospheric current flow, and then net upward field-aligned current where the current leaves the ionosphere. The global current system is also known to have continuous activity in current systems whose ionospheric paths are aligned nearly north-south and whose physical existence is shown through satellite passes through the associated field-aligned currents [Iijima and Potemra, 1978]. These systems, discussed in section 1.d, generally show considerably less ground effect than do the electrojets or substorm current wedge, for reasons illustrated in Chapter 4. Nevertheless, parameters may be added to the model to represent them, and those parameters optimized. Knowing beforehand that the perturbations associated with this near-solenoidal system may be small, one might expect the problem of finding their parameters to be somewhat ill-posed if based only on surface measurements. For that reason, the examination of satellite data which may better fix those parameters is desirable.

b. Forward Model Choice

The basic possible configurations of current systems may be classified by the flow direction of electric current in the ionosphere. It is now known that auroral zone currents are for the most part due to currents entering the ionosphere, flowing through it over some distance, and then bleeding out into space once more through field-aligned current flow. The specification of entry and exit points is tantamount to specifying the ionospheric current flow direction, provided that this current flow is basically rectilinear. The ionospheric electric current flow direction, considered as a vector in spherical coordinates based on the tilted dipole poles, may be decomposed into a north-south (meridional or poloidal) and an east-west (toroidal) component. This natural geometric division of types of current flow also corresponds roughly to expected flow directions of Pedersen and Hall currents in the observed auroral zone electric potential distribution (see section 2.c.2).

It is observed that east-west currents in the auroral zone may extend coherently across several time zones (tens of degrees), while the auroral zone itself has a north-south extent which is usually less than this, rarely more than ten degrees. Thus a normal parametrization of these currents is through a system of large east-west extent relative to its width. Since in fact the currents flow within an auroral oval having varying latitude at each meridian, in practice any requirement that these current systems have strictly east-west flow is relaxed. For a system spanning several time zones yet restricted to the auroral oval, flow will necessarily be roughly east-west, but that is all. There may be some meridional component to the roughly toroidal 'Hall' electrojet and this meridional component will be well defined. primarily by the maxima in the magnetic horizontal component underlying the electrojet and the maxima in magnitude of the vertical component near its boundaries.

North-south current flow in the ionosphere may be regarded as having closure currents at or near the borders of the auroral zone. A parametrization involving *strictly* north-south flow is not overly restrictive and in addition forces this type of current system to be different from the east-west type. The requirement for north-south flow in this sense imposes a rough orthogonality condition - there really are two types of current systems present. While in principle the less restricted east-west systems allow choices of parameters which would allow them to become nearly north-south oriented, the data usually furnish no 'motivation' for them to do so. The north-south systems are not allowed to have any toroidal current component, but have flow restricted to meridians. Since these poloidal systems, if long, contribute little perturbation even for relatively large currents, their parameters could be expected to be unresponsive to the observations. If allowed to have an unrestricted toroidal component, such current systems could rearrange themselves to have primarily toroidal current flow since such flow is most effective at producing ground perturbations and thus can be changed to efficiently get a good match to the data. This would change the originally north-south aligned currents into primarily east-west currents. In practice the problem described does occur if initially poloidal systems are not constrained to remain poloidal by not allowing any toroidal flow. By contrast, the east-west systems are constrained by ground observations of the X_m and Z components, to which they primarily respond, to cover large extents east-west. North-south systems are not well constrained from the ground but primarily respond to the Y_m component. As was the case in the discovery of the region 1 and 2 field-aligned currents which was through satellite observations, the primary constraint on the north-south system is through observations above the ionosphere. Such observations are not directly included in the present version of the Automated Forward Modelling routine. As seen near the end of Chapter 4, polar cap observations may provide some basis for solving for the poloidal currents. However, it does not harm to stress once more that inclusion of the north-south currents, if constrained only by ground observations, must be very carefully monitored.

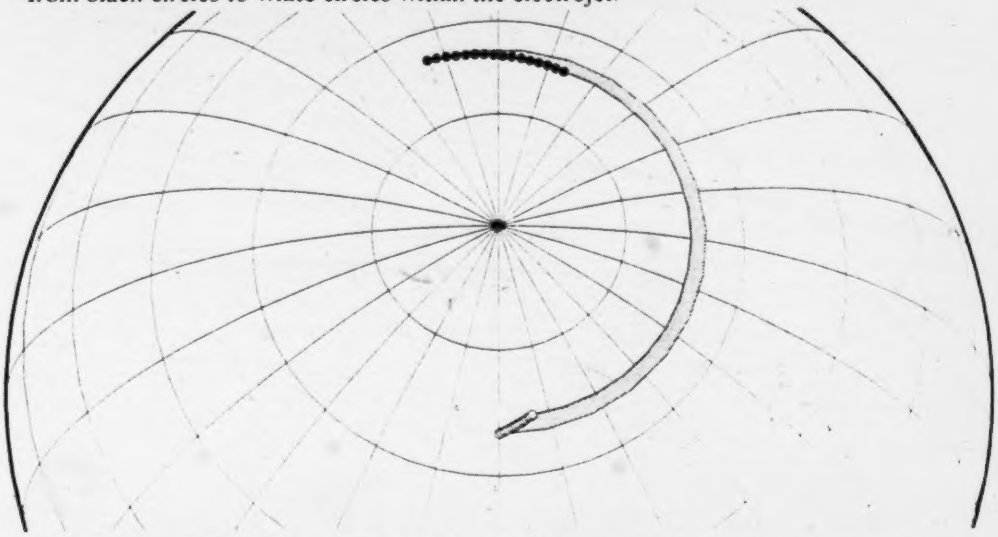
In section 4.d, representative magnetic fields arising from the two basic types of current systems were discussed. Those results parallel those of Kisabeth [1972], but are for systems exceeding the lengths which he studied. The original Kisabeth method breaks down numerically for systems of the length shown, as the spread-out point arrangement of current sources in Kisabeth's implementation of the Biot-Savart integral, using Gaussian integration, becomes dominant. A modified method was used in this study. This method consists of a simple rearrangement of the current elements used in Gaussian integration (section 4.d.1) so that there is a dense placement of elements near any given observation station. This eliminates the problems of Kisabeth's method as applied to long systems at the cost of doubling computation time. Here we briefly discuss the parametrization of the toroidal and poloidal systems. The parameters and the relationships between them indicated here are those used in the modelling.

The paths followed by currents and those defined by the intersections of sheets of field-aligned current with the ionosphere are here, for simplicity, regarded to be 'straight', which here has the following definition. A 'straight' line on the curved surface of the Earth is regarded as a locus of points along which the derivative of the latitude of points along the locus with respect to their longitude is a constant. This is a fairly natural way of parametrizing lines of current flow, but there could be others defined, and Gaussian integration can be redefined to use any parameter. Auroral currents on the largest scale generally flow parallel to the auroral oval, and thus more complex paths can be defined [e.g. Kisabeth 1972]. Here any attempt to represent such a more complex path is made using shorter 'straight' paths, appropriately joined. An example may be found in Chapter 6 where there appears to be a northward 'spur' of electrojet, deviating from what is likely the main electrojet following the auroral oval. These two systems are constrained to intersect and form one joined system in a geometrical sense, although differing currents may flow in each. It was also found necessary, in Chapter 6, to 'bend' the westward electrojet. This was done by having two systems constrained to join, but with the current also constrained to be the same in each.

1. Parameters of East-West (Toroidal) Systems

As indicated, the toroidal systems are generally of great east-west extent compared to their north-south extent. Also mentioned was the fact that it is in practice not necessary for practical reasons, nor in accord with observations, to restrict the current flow in these electrojet systems to be strictly along lines of constant magnetic latitude. In order to specify an east-west system of finite north-south extent, nine parameters are needed. The system has a region of downward current at one extremity, with two points (each consisting of a pair of longitude, latitude parameters) specifying each end of the sheet. Current then flows through the ionosphere to another current sheet, where it flows upward. As each of the two sheets requires four parameters for its general specification, there are eight geometric parameters associated with the system. The current flowing throughout the system is the ninth parameter. Figure 5.1 illustrates this type of system, with a rather distributed current 'feeding' the system and a rather concentrated current 'draining' it. Although the regions of field-aligned current flow are modelled by sheets of current, the effects seen at a distance would be similar to those of distributed regions of current. In most cases there are not enough data to allow these two cases (sheet or distributed region) to be distinguished, so that it would not be justified to add more parameters to specify more detail than in this simplest representation by sheets of field-aligned current. The proof of this statement is to be found in the generally good fits available through these simple systems as evidenced in Chapters 6 to 10.

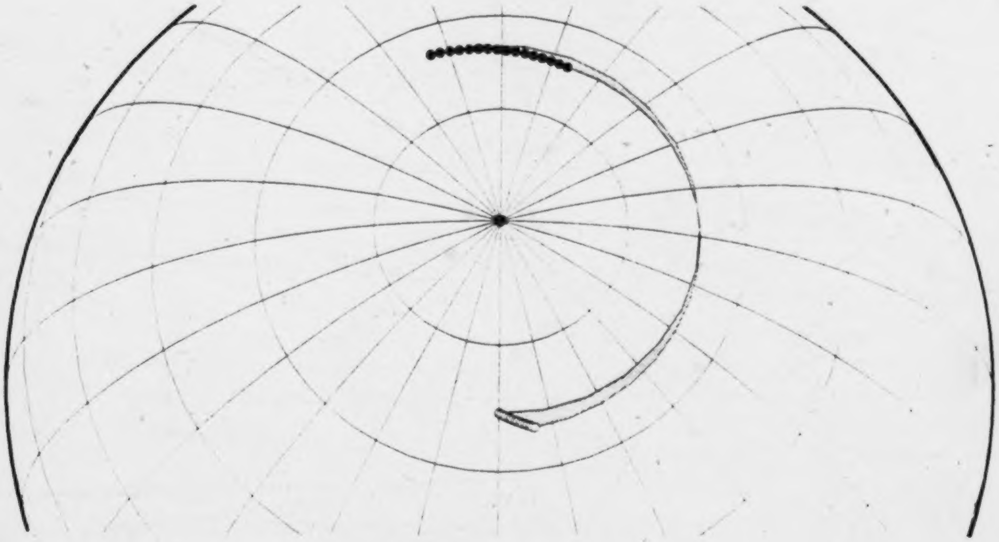
Figure 5.1 Single east-west current system (electrojet) illustrating its parameters. Downward current is indicated by black circles arrayed along a line. Region of current flow is shaded grey in proportion to total current in the system. Upward current is indicated by white circles arrayed along a line. Direction of ionospheric current flow is from black circles to white circles within the electrojet.



The optimization procedure will generally produce meaningful results if long east-west systems are used. The maxima in magnetic perturbations near or bordering the long electrojets which are normally observed in the auroral zone ensure this. In practice, all parameters within such a system may be varied and it could be expected that their optimized values would be meaningful and unique. This is to a large extent true, but certain parameters may be regarded as 'degenerate'. The first, and innocuous, form of such degeneracy may be seen, in that reversal of the northernmost and southernmost points of a current sheet does not in any way change the sheet. It does not matter which is specified first in a list of parameters. Within a current system specified by two sheets, interchanging the northernmost and southernmost points of *both* sheets forming the ends of the system, does not affect the system. A simple example is that of, say, westward current flow bounded by two lines of constant magnetic latitude. It does not matter which is specified first, the northern, or the southern, boundary. A similar and only slightly less innocuous degeneracy is that of direction of current flow. Within an east-west system, a positive sign may be attached to current flow in one direction. With some convention on order of parameters or graphical representation to represent upward or downward field-aligned flow at the ends of the system, reversal of the sign of the current causes the roles of 'upward' and 'downward' to be reversed. In most of the graphical representations in the chapters presenting results, this possible sign reversal has been corrected for, and the regions of up and down current indeed correspond to those directions. In some cases, a convention of positive sign for a direction of current flow has been adopted, and sometimes a negative value is used to indicate flow in the opposite direction. For example, in a case where positive current flow is defined to be toward the west, negative values mean that the current flow was towards the east. Generally little confusion arises from these conventions

A more troublesome type of degeneracy has a possible physical basis yet results in current systems which 'look' unphysical. As illustrated by Figure 5.2, if the end point parameters of only *one* of the bounding currents are reversed, then the joining of the end points, done according to order in a parameter list, causes crossing of the lines bounding the edges of the region having ionospheric current flow. This results in, rather than a simple quadrilateral, a 'bowtie' shape for the region of current flow. Clearly, not much of this configuration corresponds to realistic ionospheric current flow. The current density at the crossover point is in principle infinite and it is difficult to see that any naturally occurring current system would evolve to look like this. Yet such a system is indeed acceptable under the 'rules' envisaged above, which call for specification of a current system through two lines which are the loci of field-aligned current, and their joining with ionospheric current flow. Clearly these 'rules' need to be modified, and they are in practice, simply by adding the further stipulation that if the boundary lines for the ionospheric current flow cross within the longitude range of the current system, that system is not allowed.

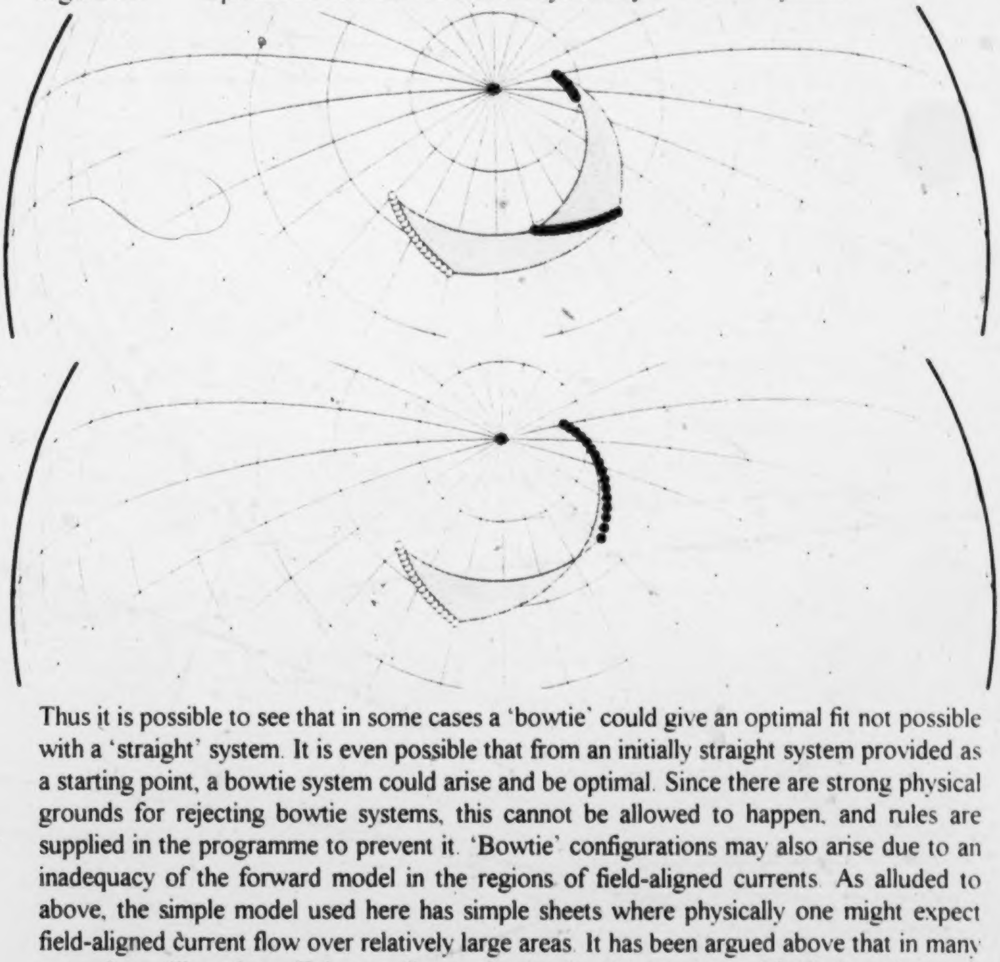
Figure 5.2 Reversal of a single bounding current sheet to result in a 'bowtie'.



While it is easy to say that 'bowtie' systems are not 'allowed' and to exclude them from the realm of acceptable results from the programme, there may be cases in which they do in some sense represent what the currents are doing. In particular, one can imagine a long system, but one not well constrained by stations all along its length. Then it is quite possible that the region of enhanced current density near the crossover point would in fact not be near a station, and thus not cause too much damage to the optimal fit to the data. Further, as opposed to rectilinear flow in the sense of 'constant derivative of latitude versus longitude' along the current flow lines, current flow generally follows the auroral oval, which is a curved path.

If there are stations straddling the auroral oval along two widely separated meridians (see Figure 5.3), then the local current flow across these may be best locally represented by straight line currents across each meridian. Due to the curvature of the auroral oval, these straight line segments may cross each meridian at a different angle. It is possible to imagine bowtie configurations which would best represent the current flow as seen at each meridian since, with arms of unequal length, maximal current flow at different angles may be had in a 'bowtie', as illustrated in Figure 5.3. This is not possible with a simple quadrilateral figure

Figure 5.3 Optimal fit of a curved current system by a 'bowtie' system.

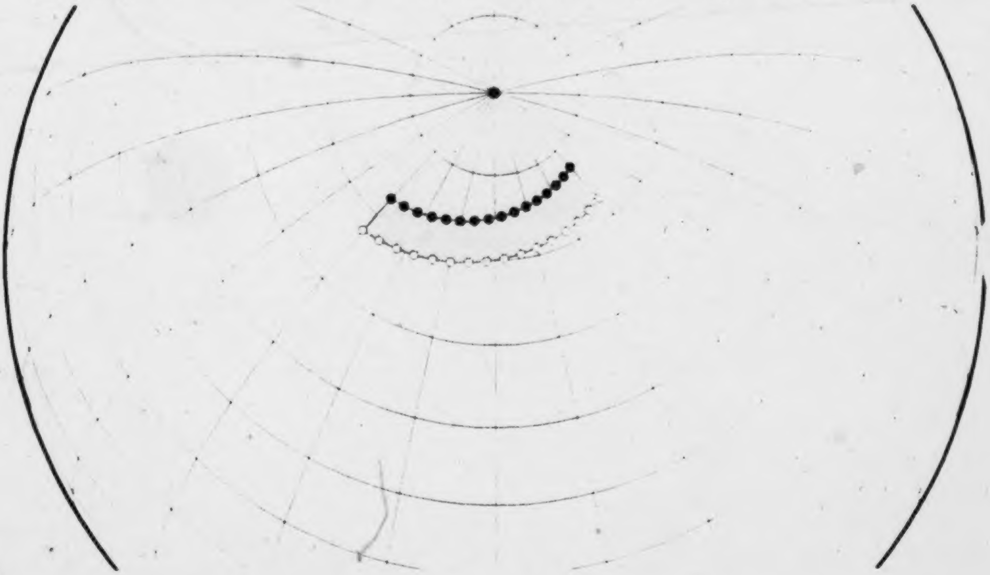


Thus it is possible to see that in some cases a 'bowtie' could give an optimal fit not possible with a 'straight' system. It is even possible that from an initially straight system provided as a starting point, a bowtie system could arise and be optimal. Since there are strong physical grounds for rejecting bowtie systems, this cannot be allowed to happen, and rules are supplied in the programme to prevent it. 'Bowtie' configurations may also arise due to an inadequacy of the forward model in the regions of field-aligned currents. As alluded to above, the simple model used here has simple sheets where physically one might expect field-aligned current flow over relatively large areas. It has been argued above that in many cases data will not be sufficient to discern the details of the regions of field-aligned current flow in any case. This in turn implies that if we are able to make models with good fits (which we are), based on the current sheet approximation, then the models must not be very sensitive to the parameters involved in specifying sheet orientation. Such indeterminacy means that the difference between a 'bowtie' system and a preferred rectilinear system may be small in terms optimality of fit, and it is even possible that the 'bowtie' system will have a lower value of an optimizing parameter (see discussion in next section). Since the routine efficiently finds the optimal fit, it can prefer 'bowties' over rectilinear systems if this is the case. However, since the problem arises precisely because of a lack of sensitivity to the parameters determining whether a system is a 'bowtie' or not, one can repair 'bowties' simply by detecting them and changing these parameters. In practice the 'untying' is done by interchanging longitudes of the ends of one sheet, without interchange of the corresponding latitudes. This untwists the 'bowtie' while retaining the longitudinal range of the field-aligned current sheet, and has proven a very satisfactory approach to resolving the problem.

2. Parameters of North-South (Poloidal) Systems

Much like an east-west system, a north-south system nominally requires eight geometric parameters to specify its four corners. However, due to the ill-posed nature of the problem of determining the parameters of north-south systems from the ground, it is necessary to restrict the current flow in these systems to be strictly along meridians (i.e. purely poloidal). This effectively removes two parameters from play, meaning that only seven parameters are needed. These are the longitude of the most westerly points, the latitudes of up and down current at that longitude, the corresponding three parameters at the most easterly point, and the current. A north-south system is illustrated in Figure 5.4.

Figure 5.4 North-south current system as specified by seven parameters



In practice, the north-south system is not shown in illustrations in this thesis, even where incorporated into the modelling. The practical reason for this is that the east-west systems have by far the largest ground effect and thus must be illustrated, whereas the north-south system has little ground effect and would simply clutter the map. Another aspect is that not only is the north-south system generally acknowledged to be in essentially the same place as the east-west electrojets [Senior and Robinson, 1982], but these parameters for this system would be poorly defined by ground modelling in the first place. Since this is the case, in most cases there has been a relationship established between the east-west and north-south systems (usually that their borders are essentially identical), and the parameters of the north-south system have not been independently derived in modelling, apart from (usually) the total current in the system.

3. Constraints on Parameters

Various parameters may not be entirely independent or may have known properties which are not part of the model. In this case it may be useful to constrain the parameters with respect to one another or an external value. The following discussion explains how this is in practice realized in this project. In other methods which have been discussed(in Chapter 4) the issue of constraint may also arise. For AMIE, some aspects are considered in Appendix I, Section III of Knipp [1989], and use of the statistical expected value matrix **C**, as a constraint was discussed in Section 4.c above, so constraint in other methods is not discussed further here.

A constraint may be implemented in Automated Forward Modelling in at least two ways, both of which have in fact been used. Since in many cases a constraint can be expressed as a non-linear equation or operation, it can be accommodated in a nonlinear fitting routine in a natural way by simply incorporating the rule of constraint into the calculation routine of the forward model. In this way, the forward model may contain rules (the constraints) about the interrelation of its parameters. The second, less rigid, method of enforcing a constraint is to impose a penalty on an optimization variable (see next section) if the constraint condition is not met. This is in principle similar to the approach taken in AMIE, although here it was generally applied in a multiplicative fashion rather than the basically additive fashion employed in AMIE.

An operational example may be offered as follows. Consider a case in which the latitudinal width of an electrojet has been well determined along one meridian where there is a magnetometer chain. Global modelling is to be done, including an adjacent region (a gap) where there are insufficient stations to well constrain the width of the electrojet. Based on past knowledge, or better yet based on some possibly available source of ancillary information such as a satellite image, it may be reasonable to constrain the electrojet latitudinal width to be constant. Incorporating the constraint rule into the forward modelling could be done simply by requiring the parameters specifying the latitudes of the ends of the upward and downward current sheets to be separated by the same amount. In practice, rather than solving for the latitudes of the four corners of the region bounding ionospheric current flow, one would search for the longitudes and for one latitude (say the most southerly) at each end, the other latitude at each being calculated through addition of the latitudinal width. In this case one effectively removes two parameters from the system, since the sheets of field aligned current at the ends of the region of ionospheric current flow have one less free parameter each. The alternate approach is to incorporate the constraint rule into the optimization variable which is systematically reduced in optimizing. In this case, one would increase that variable based on how far the difference of latitudes in a proposed forward model deviated from the desired value. Since the optimization involves reduction of the variable, increasing it as a proposed solution deviates from some favored solution will result in a tendency for the solution to remain near the favored solution unless strongly pulled away by the data. This latter approach is less rigid since the data could dictate violation of the constraint if sufficiently 'constraining'. For example, if one later got another set of meridian chain data to fill the gap referred to above, the first method mentioned would not respond to the new information this would give about electrojet width, whereas the second method would. The changes in the final result would vary to a degree determined by the relative weighting of the constraint condition and data.

In practice the difficulty in implementing constraints is largely one of programming. In this project a simple parser was developed to operate on an auxiliary file which told which parameters were to be varied and which to be held constrained. In the case of a parameter to be constrained, it was possible to specify that its value was to have some relationship to that of some other parameter, or to be held fixed. Simple operations involving another parameter and a constant were implemented: for example it was possible to constrain a northern corner of the current flow quadrilateral to be some latitudinal separation from the corresponding southern corner. An example will now be given to illustrate use of constraints in modelling. Table 5.1 shows parameters used in an actual model run to specify initial conditions for two currents. As mentioned above, the quadrilateral region within which current flows is specified by eight geometric parameters and by the total current flowing within the region. The geometric parameters may be split into those associated with the sheet of upward 'drain' current from the ionosphere and those of the downward 'feed' current into the ionosphere. These sheets are in turn represented by their extreme points, each specified by a longitude (negative for west longitudes) and a latitude. The set of four such pairs of points, plus the current, completely specifies a three dimensional current system as used in modelling here. Within Table 5.1, the parameters for the upward current sheet in the second system are identical to those for the downward current sheet in the first system. This is simply a way of tying the two systems together: since the current in the second, more westerly system is greater than that in the first system, the net effect is that current flows into the ionosphere at this junction.

Table 5.1 Current Parameters for Two Current Modelling Run

Longitude Up 1	Latitude. Up 1	Longitude. Up 2	Latitude. Up 2	Longitude Down 1	Latitude Down 1	Longitude Down 2	Latitude Down 2	Current (MA)
-104.528	62.3278	-10.2782	72.2485	-26.4025	63.4717	41.7283	77.6546	0.196215
-131.174	73.8745	-190.445	76.2815	-104.528	62.3278	-10.2782	72.2485	0.164395

Table 5.2 Rules for Variation and Constraint of Parameters for Two Currents

Longitude Up 1	Latitude Up 1	Longitude. Up 2	Latitude. Up 2	Longitude Down 1	Latitude. Down 1	Longitude Down 2	Latitude Down 2	Current (MA)
1	1	1	1	1	1	1	1	1
1	1	1	1	-1.1	-1.2	-1.3	-1.4	1

Table 5.2 shows the parameters which form rules for the variation of the corresponding physical parameters in Table 5.1. In this case these specify that throughout the modelling run, the parameters corresponding to a '1' are allowed to freely vary. This applies to all parameters of the first system, and to those of the upward current sheet and current magnitude of the second system. The parameters of the downward current for the second system have negative numbers associated with them. A negative number indicates that a constraint is to be applied. In this case, the rule is simply that a parameter will be made identical to a parameter within another current system. The number of that system (1, 2, etc.) must be given as the integral part of the rule parameter, while the number of the parameter within that system to use is given by the decimal. For example, the rule parameter for 'Latitude Down 1' within the second system is -1.2. This indicates that the latitude of the first point of the downward current in system 2 is to be numerically equal to the value of the second parameter from system 1. That may be seen to be the latitude of upward current from system 1. In this way the two systems are required to be joined in the manner described in the preceding paragraph, at all times during convergence of the model. If the rule parameters are not specified, they default to 1, allowing free variation of all physical parameters. In that case, an initial system as specified in Table 5.1, with joined currents, could be broken into two unlinked current loops if that would optimize the fit to the data. The number of effective parameters, if all are given free variation, is larger than in the constrained case, and with more freedom of variation, it is quite possible for a better fit to be obtained in such circumstances. The choice to enforce a constraint is motivated by the desire to have a physically realistic system, possibly at the expense of the goodness of fit. Of course, judgment would have to be used. If a much better fit is obtained without constraint, the possibility that the constraint is inappropriate must be considered.

The parsing language allows other constraints to be imposed at run time. Should it be desired to hold a physical parameter fixed at its input value, a '0' may be used as the corresponding rule parameter. For example, if the '1' for the current in the first system in Table 5.2 were changed to a '0', the input value of the current (0.196215 MA) would be used throughout the modelling run, and the optimization would be done based on only the geometric parameters. It was mentioned above that constraints could be introduced to fix the relationship among various parameters, beyond simply making them identical. The example mentioned was that it would be possible to specify a width for the electrojet. To make this clearer, consider the Latitude Up 2 and Latitude Down 2 parameters in Table 5.1. If the corresponding rule parameters in Table 5.2 were changed from '1' (for free variation) to -1.2+5, and -1.5+5, respectively, then the Latitude Up 2 and Latitude Down 2 physical parameters would be constrained to be 5 degrees larger than Latitude Up 1 and Latitude Down 1. In practice, this type of more complex restraint was rarely used, while nonvariation or linking of parameters were used extensively. In practice, use of the simple parsing language proved very effective in constructing joined systems or studying variations based on a limited set of parameters.

A very important application of constraint was in tying the north-south systems, with little ground signal, to east-west systems having sufficient signal to allow solution. In this case the geometric parameters of the systems were related to each other so that those of the east-west system were those of the north-south, and the current in the north-south system freely solved for. It should be noted that this does not mean that the position of north-south system current flow was removed from the system and simply not solved for. Rather, the combined fields of both systems would be matched to determine the geometric parameters, which are constrained to be the same for both systems. Normally those parameters would be more heavily constrained by ground observations due to the east-west part of the total current flow, as these are larger.

c. Optimization

Before considering in detail the optimization by nonlinear least squares which is actually used in Automated Forward Modelling, a brief discussion of optimization theory will be presented. In this way it will be seen that there are specific reasons for the choice of that method and for the Levenberg-Marquardt implementation used.

Rice [1983] discusses three ways in which approximation problems arise. One classical example is in the approximation of mathematical functions. Although such approximations are at the basis of the numeric computation required in digital computing, and thus fundamental in that sense, this is not the approach relevant here. Similarly, smoothing and analysis of data, his third class, may be used, but in his sense there is little 'specialized information' available in this case. A more appropriate approach is found in his class of problems involving 'representation and compactification' of data. Here 'specialized information' is available which allows a fitting model to be proposed. Parameters within the model may be varied to obtain an optimal model. This is the approach which is taken here and referred to as 'forward modelling'. But as opposed to merely the numerical analyst's interest in 'compactification', we ascribe a physical meaning to a model which fits well. By proposing a forward model in which the parameters correspond to physically measurable quantities, we claim that the observed data resulted from a physical system having characteristics analogous to those described by the best-fit model parameters.

Optimization attempts to deliver those best-fit parameters given the proposed model and the data. 'Closeness' of the model to the data is measured by the *norm* [Rice, 1983], of which many may be defined, and of which one must be chosen. To illustrate, consider the approximation of a set of N points of data, y_i, where i varies from 1 to N, by the function $F(\mathbf{a},x_i)$, or simply $F(\mathbf{a})$, where **a** is the set of parameters appropriate to the model represented by F and the x_i are ordinates corresponding to the data values. The deviation of model from observation, with notation $\|\mathbf{y} - F(\mathbf{a})\|$, is called the *norm*. The maximum deviation or Chebyshev norm is simply $\|\mathbf{y} - F(\mathbf{a})\|_\infty = \max|y_i - F(\mathbf{a}, x_i)|$ for $1 \le i \le N$. The sensitivity of this norm to the single largest deviating point in the data set is itself a disadvantage and it is not easy to minimize, largely due to the discrete dependence on whichever point that happens to be. Thus it is not suitable for our type of problem. More responsive to the entire data set is the norm representing average error, the least deviation or **L**$_1$ norm, defined as $\|\mathbf{y} - F(\mathbf{a})\|_1 = \sum_{i=1}^{N}|y_i - F(\mathbf{a}, x_i)|$. This norm is less responsive to outlier points, as is indicated by its discussion in 'Numerical Recipes' [Press *et al.*, 1992] under the heading of 'robust statistics'. That it is also difficult to use in practice is indicated by the fact that the sole example of its use given there is for fitting straight lines. In practice the easiest norm to use is the least-squares or **L**$_2$ norm, defined by

$$\|\mathbf{y} - F(\mathbf{a})\|_2 = \sqrt{\sum_{i=1}^{N}\left(y_i - F(\mathbf{a}, x_i)\right)^2}.$$

This norm represents the distance between fit and data in an N-dimensional space, and it is clear that lessening the deviation of the fit function from the points reduces the value of the norm. For both the **L**$_1$ and **L**$_2$ norms, it is possible to introduce weighting functions, which may be chosen so as to make the norms for discrete data sets resemble integrals which would arise through trapezoidal integration in continuous data, or may reflect the relative reliability of each data point. In practice the weights enter as a multiplicative factor (say w_i) in each term of the sum used to form the norm, or as error weight terms inversely proportional to some error value σ_i. This leads naturally to the consideration of the effects of errors in the data in fitting and to a more rigorous reason to prefer a form of weighted **L**$_2$ norm. Following Press *et al.* [1992], we consider that the data points y_i are normally distributed about what the 'true' model would give (with this distribution being due to random errors). The probability of obtaining the measured data set of N points is then the product of probability of each point comprising it:

$$P \propto \prod_{i=1}^{N}\left\{\exp\left[-\frac{1}{2}\left(\frac{y_i - F(\mathbf{a}, x_i)}{\sigma_i}\right)^2\right]\Delta y\right\}$$

where Δy is an arbitrary constant reflecting a small interval in y within which the probability density of y_i is evaluated, and σ_i is the standard deviation of the distribution of data points. Intuitively (and we are reminded in 'Numerical Recipes' that statistics is *not* a branch of mathematics) one must identify that set of parameters, considered as variables, which would maximize the probability of the data set (which is given and not to be changed) to be that represented by the parameters.

Minimizing the negative of the logarithm of P will maximize P, and that will require the quantity chi-squared (χ^2),

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - F(\mathbf{a}, x_i)}{\sigma_i} \right)^2$$

to be minimized. This approach to justifying the use of χ^2 minimization for maximum likelihood estimation of parameters is based on the assumed normal distribution of data errors. Even if that is not the case, the inverse errors may be considered as weights, resulting in those points with the smallest errors receiving the highest weighting. It is also possible that the errors are unknown, in which case they may be set to unity and χ^2 minimization is equivalent to seeking the minimal $\mathbf{L_2}$ norm.

In the foregoing there has been no discussion of the functional dependence of $F(\mathbf{a},x)$ on either the parameters $\mathbf{a}$ or the independent variable x. In practice, methods of solution of the optimization problem depend on this functional form. Methods in which the dependence on model parameters is linear are, not surprisingly, relatively amenable to direct solution, while a more general dependence requires a more involved approach. We will consider first how the linear case may be applied to magnetic data. The form of solution presented will be recognized as being the basic form used in the AMIE technique discussed in Section 4.c.

Assuming a linear dependence on the model parameters $\mathbf{a}$ does not imply any particular functional dependence on x, that dependence entering in the M basis functions X(x) put forward as part of the proposed model. Since the following discussion applies equally to discrete and continuous ordinates, the subscript on x is sometimes omitted for convenience. M should be chosen so that there are more data points than parameters being solved for if one expects a unique best solution; in other words $N \geq M$. The quantity $\nu = N-M$ is known as the number of degrees of freedom in the problem, and it should be positive (note that ν is not to be confused with the error vector $\boldsymbol{\nu}$ discussed previously in connection with AMIE). The model's form in this linear case is

$$F(\mathbf{a}, x) = \sum_{k=1}^{M} a_k X_k(x)$$

and the χ^2 becomes

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - \sum_{k=1}^{M} a_k X_k(x_i)}{\sigma_i} \right)^2$$

where it is important to note that the basis functions appear evaluated at each data point abscissa. In this equation, it is the parameters a_k which are variables, the set of (x_i, y_i)

observation pairs being preselected, and the basis functions $X_k(x)$ by hypothesis being independent of the parameters a_k. It is thus very clear that minimization takes place in the space of the parameters. The solution of this optimization problem is called a *linear* least squares solution: it is linear because the dependence on each parameter is linear. The dependence of the basis functions on the independent variable(s) may well be nonlinear. That being the case, this approach is widely used in the representation of magnetic data by functional expansions, such as in equivalent current fitting by J. K. Walker *et al.* [1995], or as an important part of the AMIE procedure (see Chapter 4). It is instructive to examine how the optimal parameters are determined in linear least squares fitting, for comparison with the nonlinear method used in this work. Letting **A** be the matrix with elements $A_{ij} = \frac{X_j(x_i)}{\sigma_i}$ (the design matrix), and a weighted observation vector **b**, where $b_i = \frac{y_i}{\sigma_i}$, the expression above becomes $\chi^2 = \sum_{i=1}^{N}(b_i - \sum_{k=1}^{M} a_k A_{ik})^2 = (\mathbf{b} - \mathbf{Aa})^T(\mathbf{b} - \mathbf{Aa})$, leaving the vector **a** of coefficients (parameters) to solve for. This solution may be found by making the derivative of χ^2 with respect to each parameter be zero. Taking the set of such derivatives, one obtains the M normal equations (k=1,...,M) to be solved simultaneously:

$$0 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}\left[y_i - \sum_{k=1}^{M} a_j X_j(x_i) \right] X_k(x_i) = (\mathbf{A}^T \cdot \mathbf{A}) \cdot \mathbf{a} - \mathbf{A}^T \cdot \mathbf{b}$$

To obtain a solution, a matrix must be inverted and there are several standard ways to obtain such an inverse. The matrix form may be rearranged and symbolically solved to give a compact expression of the solution, which is $\mathbf{a}=(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{b}\mathbf{A}^T$. Since all variations in the parameter space are linear in the parameters, the solution is particularly simple.

Linear inversion is suitable in cases for which a sum of weighting factors multiplying basis functions allows good representation of data. The basis functions need not be orthogonal: in a polynomial expansion, for example, they would not be. In many spaces an expansion in orthogonal functions is possible; for near-Earth magnetic fields these expansions could be in spherical harmonics as in the IGRF or in some modified spherical harmonic expansion as has been detailed in Chapter 4 in discussing the KRM and AMIE techniques. As also discussed there, such parameters in general do not have direct physical meaning, which is considered here as a disadvantage. Referring to Section 5.a, we might also note that the linear expansion needed for an eccentric dipole is very inefficient and that a more physical model could have considerably fewer parameters. However, the dependence of the field on parameters specifying the position of the dipole, as may be seen by referring to the expressions for dipole fields in Chapter 1, would be highly nonlinear. Only the parameter specifying the strength of the dipole would affect the observed fields in a linear fashion. The cost of nonlinearity in the model is that a simple linear fitting approach no longer is applicable.

Where the data have linear dependence on the parameters, χ^2 has quadratic dependence on them. Requiring its derivative to be zero restores linearity since the derivative of a quadratic is a linear function. This yielded an algebraically soluble (by matrix inversion) set of equations. One can equally do the same steps for arbitrary dependence of the model on its parameters. The problem arises at the last step since the resulting analogue of the normal equations will generally not be soluble. For this reason, most attempts to solve optimization problems with nonlinear dependence on parameters are done by following gradients of χ^2 in the parameter space. Usually this must be done numerically. Some numerical methods popular about 1970 are discussed in the reprint of the book *Numerical Methods that Work* by Acton [1990] and in the second edition of Hamming's [1973] well-known treatise. The Fletcher-Powell method and its variants appear to have been most popular at that time. The related Levenberg-Marquardt method is now the most widely used, since it 'works very well in practice and has become the standard' [Press *et al.*, 1992, p. 683]. Both are variable metric methods which attempt to find an optimal step size in parameter space, and both are related to Newton's method in the sense of following gradients toward a solution. Here the Levenberg-Marquardt algorithm is described largely through putting together and generalizing material contained in various places in *Numerical Recipes in C* [Press *et al.*, 1992]. The fact that the algorithm is the present-day 'technique of choice' and more references, including several to its widespread commercial implementations, are found in a recent summary [Lampton, 1997]. A concise description of the routine's operation, more coherent than that in *Numerical Recipes* is also given there, along with a test which finds a strategy permitting more computational efficiency.

We assume that χ^2 for a forward model which has parameters can be represented in the space of those parameters as a differentiable function. This is the case for our magnetic field models if using physically meaningful parameters. Then the χ^2 function can be expanded about any point $\mathbf{a_0}$ in M-dimensional (as above) parameter space, to have a value at another point $\mathbf{a}$ (i.e. for a different set of parameters) which is

$$\chi^2(\mathbf{a}) = \chi^2(\mathbf{a}_0) + \sum_{i=1}^{M} \frac{\partial \chi^2}{\partial a_i}\Big|_{a_0} (a_i - a_{0i}) + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \frac{\partial^2 \chi^2}{\partial a_i \partial a_j}\Big|_{a_0} (a_i - a_{0i})(a_j - a_{0j}) + \dots$$

This may be truncated within the domain of convergence and sufficiently close to $\mathbf{a}$, to be written more compactly in a vector form as

$$\chi^2(\boldsymbol{a}) \approx \chi^2(\boldsymbol{a}_0) - \boldsymbol{b}\cdot(\boldsymbol{a} - \boldsymbol{a}_0) + \tfrac{1}{2}(\boldsymbol{a} - \boldsymbol{a}_0)\boldsymbol{A}(\boldsymbol{a} - \boldsymbol{a}_0).$$

where $b_i = -\frac{\partial \chi^2}{\partial a_i}\Big|_{\mathbf{a}_0} = -(\nabla \chi^2)_i\Big|_{\mathbf{a}_0}$ is the negative gradient vector of χ^2 in parameter space, and $A_{ij} = \frac{\partial^2 \chi^2}{\partial a_i \partial a_j}\Big|_{\mathbf{a}_0}$ is the so-called Hessian matrix, both evaluated at the point of expansion. This expansion is a quadratic form. It is well known that near an extremum (where the gradient is zero) the quadratic form is the lowest non-trivial approximation to a function. We can see the family resemblance of the Levenberg-Marquardt method to Newton's method by the following considerations. From the expansion, the general form of the gradient of χ^2 can be readily calculated as $\nabla\chi^2(\boldsymbol{a}) = \boldsymbol{A}(\boldsymbol{a}-\boldsymbol{a}_0)-\boldsymbol{b}$. Newton's method could be used to find the extremum in a seemingly straightforward way. If the gradient is set to zero and the point **a** thus regarded as the extremum point, one simply has $\boldsymbol{A}(\boldsymbol{a}-\boldsymbol{a}_0)=\boldsymbol{b}$, with solution $\boldsymbol{a}=\boldsymbol{a}_0+\boldsymbol{A}^{-1}\boldsymbol{b}$. As with Newton's method in one dimension, this solution is to be regarded in an iterative sense, since the Hessian matrix and the gradient vector **b** are both evaluated at the expansion point and may not be representative of the behaviour of the function all the way to its extremum. If the expansion terms retained are sufficient, the solution may be a better approximation to the extremum than was the original point $\mathbf{a}_0$. Despite the fact that convergence of Newton's method is slow, we are likely better off now than with the initial guess. And in fact, since a quadratic form fits well near an extremum, the approximation made above is probably now a better approximation, leading to further improvement on the next step. This might lead one to expect that Newton's method would be an adequate approach to general non-linear χ^2 optimization. However, it may be that the initial guess is not in fact close enough to the solution for a quadratic approximation to be valid. In that case, a linear function, rather than a quadratic, is the lowest-order best approximation to the function near any point of expansion. The gradient from that linear function can be followed downslope and in that manner a minimum might be expected to be approached. The problem with this approach is that one does not know *how far* to follow the gradient. It is this problem which the Levenberg-Marquardt process attempts to solve, and in the process bring in the power of quadratic fitting where appropriate.

Descent along the gradient can be examined through the linear approximation $\chi^2(\boldsymbol{a}) \approx \chi^2(\boldsymbol{a}_0)-\boldsymbol{b}\cdot(\boldsymbol{a}-\boldsymbol{a}_0)$ which is not useful in that form since **a** is unknown and there is no particular way to find it. What we can claim is that *if* we had the right elements for **a** and thus for the step $\mathbf{a}\text{-}\mathbf{a}_0$, we could go down the slope to somewhere near the minimum. We would know if we had done so since χ^2 there would be less than at the starting point. We also know that the step must be locally roughly parallel to the negative gradient, that is, to **b**. This can be written for each component with an undetermined constant which will be called (with some forethought) $\lambda\alpha_{ll}$, so that the equation for the l-th component of the step is $\lambda\alpha_{ll}(\mathbf{a}-\mathbf{a}_0)_l = \mathrm{b}_l$. Recalling that the quadratic approach led to the equation $\boldsymbol{A}(\boldsymbol{a}-\boldsymbol{a}_0)=\boldsymbol{b}$, we note some similarity of form between these two equations. The first is equivalent to a diagonal matrix multiplying the step, whereas the second multiplies by the generally non-diagonal Hessian. The dimensions of the elements of **A** and of $\lambda\alpha_{ll}$ must be the same, so if we regard λ as a dimensionless quantity, then the dimensions of α_{ll} must be those of the elements of $\boldsymbol{A}$. Without any loss of generality, we can claim that the α_{ll} could in fact *be* the diagonal elements of $\boldsymbol{A}$, and that if they are not, then λ could be adjusted so as to at least partially compensate. This rather loose approach allows the *combination* of the two equations into one, with λ to be determined. If δ is the identity matrix, we will regard the system $\boldsymbol{A}'(\boldsymbol{a}-\boldsymbol{a}_0)=\boldsymbol{b}$ as in some way containing the desired solution, with $A'_{ij}=A_{ij}(1+\lambda\delta_{ij})$ the elements of the new combined matrix A'. We must now formulate a method involving λ which allows this new system to solve the optimization problem. First we note that two extremes of value for λ allow extraction of the original two equations. For $\lambda=0$ we recover the quadratic approach, while if λ is very large, A' becomes diagonally dominant and the linear case is effectively in force. For intermediate values of λ, the new system will have a solution analogous to the Newton's method case: $a = a_0 + A'^{-1}b$. By choice of λ, one can cause this solution to respond to being near a minimum (quadratic) or on a slope far from a minimum (linear). Control over λ will depend on the acid test of getting nearer the minimum: whether χ^2 at the new **a** is smaller than that at the initially-guessed value $\mathbf{a}_0$. If, for some intermediate value of λ, χ^2 has not become smaller, then presumably one is on the slope and an increased value of λ would make the equation correspond better to the local situation. If (as desired), χ^2 has become smaller, then one is likely nearer to the minimum and a closer resemblance of the modified system to a quadratic form is required. This can be arranged by reducing the value of λ. Ultimately, when the value is reduced to near zero, the pure quadratic form appropriate to being near a minimum is used and convergence accelerates. Based on these observations, the 'recipe' for use of the Levenberg-Marquardt technique given in *Numerical Recipes in C* [Press *et al.*, 1992, p. 684] may now be cited. Their version begins optimistically with a rather low value of λ: in practice an optimal starting value must be determined by the user. The recipe is:

1. For a starting guess at the parameters $\mathbf{a}_0$, compute $\chi^2(\mathbf{a}_0)$.
2. Pick a small value of λ (they suggest 0.001 in hopes of starting near a minimum).
3. Solve $A'(a - a_0) = b$ to get $a = a_0 + A'^{-1}b$, then evaluate $\chi^2(\mathbf{a})$.
4. If $\chi^2(\mathbf{a}) > \chi^2(\mathbf{a}_0)$, one has failed to approach the minimum and must increase λ significantly (they suggest a factor of 10) and repeat step 3.
5. If $\chi^2(\mathbf{a}) < \chi^2(\mathbf{a}_0)$, one is nearer to the minimum, so **a** is a better guess than $\mathbf{a}_0$. One must replace $\mathbf{a}_0$ in step 3 by **a**, decrease λ to reflect being nearer the minimum, and redo 3.

In practice, steps 4 and 5 must also include some criterion to determine when to give up. This may be some tracking of successive values of χ^2 so that when it stops changing much, one stops (which is the method recommended in *Numerical Recipes*), or simply counting the number of iterations. The latter is a more practical method for implementation, but should be coupled with examination of output at each step so that the 'run' of χ^2 is monitored and has a reasonable behaviour (see Section 5.g). Lampton [1997] finds greatest efficiency in a simple problem when λ is added directly to diagonal elements of the Hessian matrix **A**. This is referred to as additive damping while the scheme detailed is called multiplicative damping. The multiplicative approach is suggested when the parameters have widely varying scales, and with the large number of parameters used in this study, that is likely to be the case. In both additive and multiplicative cases, Lampton cites multiplicative factors for rescaling λ of 0.1 and 10 as commonly used, although his study indicates that these should not be reciprocal factors for most efficient additive damping. In this study the upward rescaling of λ has been done with factors from 3 to 10 and reciprocals have been used when it has been necessary to decrease λ.

This is an appropriate place for a brief further discussion of the Hessian matrix $A_{ij} = \frac{\partial^2 \chi^2}{\partial a_i \partial a_j}\Big|_{\mathbf{a}}$. Since the form of χ^2 is known, as are the forward model dependencies on the parameters, there is no problem in taking successive derivatives with respect to the parameters to obtain the functional form of the Hessian elements:

$$A_{ij} = \frac{\partial \chi^2}{\partial a_i \partial a_j} = 2\sum_{k=1}^{N} \frac{1}{\sigma_k^2}\left[\frac{\partial y}{\partial a_i}\frac{\partial y}{\partial a_j} - \left(y_k - y(\mathbf{a})\right)\frac{\partial^2 y}{\partial a_i \partial a_j}\right]$$

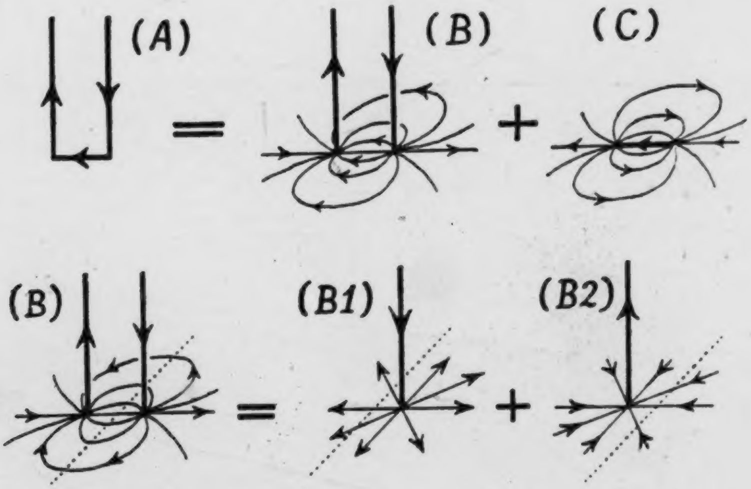
In principle, this form of the Hessian may be used with the algorithm outlined above. In practice, however, there are motivations for neglecting the second derivative terms. In the application presented here, the derivatives are calculated numerically and this is computationally expensive. This is equally the case if the minimization technique is applied with analytical derivatives [Connors, 1993]. We may be able to convince ourselves that second derivative terms should be small (this is certainly the case in linear modelling where they are zero, but in that soluble case we do not need to use a gradient following technique). Press *et al.* [1992] supply some more convincing arguments, including that the term multiplying the second derivatives is the error of the model at each point which should be randomly distributed with mean zero. They note also that second derivative terms can be destabilizing while not affecting the final result, only the route to it. So in practice the second derivative terms are rarely used in implementations of the Levenberg-Marquardt algorithm. Without them, the algorithm is referred to as a 'half-Newton' method: with second derivatives included it would be 'full-Newton'.

We now proceed to discuss uniqueness of solutions in minimization with ground magnetic data. Other practical aspects of using the Levenberg-Marquardt algorithm to solve problems involving magnetic effects due to current systems are presented in Section 5.g below.

d. Uniqueness

The historic conflict in Space Physics about the nature of the basic current systems causing ground disturbances has been alluded to. Useful insight into why non-uniqueness arises in this case has been given by Fukushima [1969, 1976], although the often-quoted 'Fukushima theorem' is misleading for reasons given below. In Fukushima's illustrative example, field lines are considered to be vertical (a good approximation in the auroral ionosphere), and the ionospheric conductivity is considered uniform and isotropic (a much less correct assumption). The Birkeland current system, a loop consisting of downward field-aligned current, an ionospheric path, and an upward field-aligned current, may have its ground magnetic effects duplicated by a system entirely in the ionosphere with the same ionospheric path and surrounding closure currents, as illustrated in Figure 5.5. Such a system would be similar to the Chapman-Vestine equivalent system previously alluded to. That these produce identical ground effects may be seen by considering the difference currents between the two cases. These amount to the field aligned currents and closure currents. A set of closure currents may be made by considering each field-aligned current to feed into the ionosphere and flow radially outward in that plane from the point of contact. Each such system produces no ground effect. They are each symmetric about the vertical axis, and this symmetry is at the heart of the 'Fukushima theorem'. The integral form of Ampère's law in this symmetric case states that $2\pi r B = \mu_0 I$, where B (in the azimuthal direction) at radius r is thus proportional to the total current (here I) inside that radius. Below the ionosphere, no current flows and thus the ground-level B from this current system is identically zero. Recalling that this current system is the difference between the Birkeland and the Chapman-Vestine systems, one sees that there are no magnetic signatures of that difference and hence no way to use sub-ionospheric measurements to distinguish the two cases. Although the example is merely illustrative and based partly on unrealistic assumptions, clearly the uniqueness of any model current system put forward must be justified by other considerations than only the ground observations. This simplified theoretical approach is entirely consistent with modelling results for meridional systems as shown in Chapter 4. In the Fukushima example it is *impossible* to determine the current flowing in the field-aligned currents from ground observations. In practice, with poloidal model current systems, the problem is simply very ill-posed and results must be interpreted with great caution and preferably in conjunction with auxiliary data.

Figure 5.5 Fukushima [1969] equivalent currents for a field-aligned current pair. System (B) is the difference current between Birkeland system (A) and Chapman system (C). (B) is equivalent to (B1) plus (B2), neither of which produce ground magnetic fields.



More recently the use of simplified models (such as those discussed in section 4.a.3) by modellers of satellite magnetic observations has led to the belief that the Region 1/2 current systems, characterized by north-south paired sheets of field-aligned current joined by ionospheric currents, are basically solenoidal in nature and do not produce significant ground effect. The simplified models produce their largest field in the region between the field-aligned sheets and may be used for a reasonable interpretation of field-aligned current based on observations there. Detailed modelling [Kisabeth, 1979], as discussed in Chapter 4, suggests that there may be ground perturbations observable from the Region 1/2 currents, particularly in the polar cap. Simplified models such as those of Fukushima or those used in analysis of spacecraft data would not suggest this to be the case. The more refined models might allow determination of the Region 1/2 current strengths from the ground, but only poor localization of these currents. This is investigated below, in several chapters, where the addition of Region 1/2 currents appears to improve model fit.

In addition to the formal non-uniqueness demonstrated by Fukushima for magnetic problems, there may be non-uniqueness associated with the fitting problem. The investigator searches for a 'best' fit indicating a close resemblance of the proposed model to reality when the appropriate parameters are chosen. It may be relatively straightforward to minimize in parameter space by finding a point where variations are zero but that does not guarantee that the fit is 'best' as there may exist other such points. Such points correspond to local minima in the parameter space. With the effective changes in length scale in the Levenberg-Marquardt technique, such local minima might be expected not to be a great problem. It has been found through experience that a poor choice of initial scale parameter can in fact lead to being caught in a local minimum. Given the large number of parameters involved in magnetic modelling, particularly on a global scale, it is hard to be absolutely certain that local minima are not hindering solution. Such uncertainty must be accepted once minimization with more than a few parameters is attempted. As stated in Numerical Recipes [Press *et al.*, 1992], "this kind of problem is generally quite difficult to solve". Quality of solution, including assessment of the likelihood of having attained a global minimum, must be quantitatively assessed from the behaviour of χ^2 throughout the solution, and from the reasonableness of the final parameters. It is similarly difficult to quantify error bounds when a large number of parameters vary.

e. Error Bounds

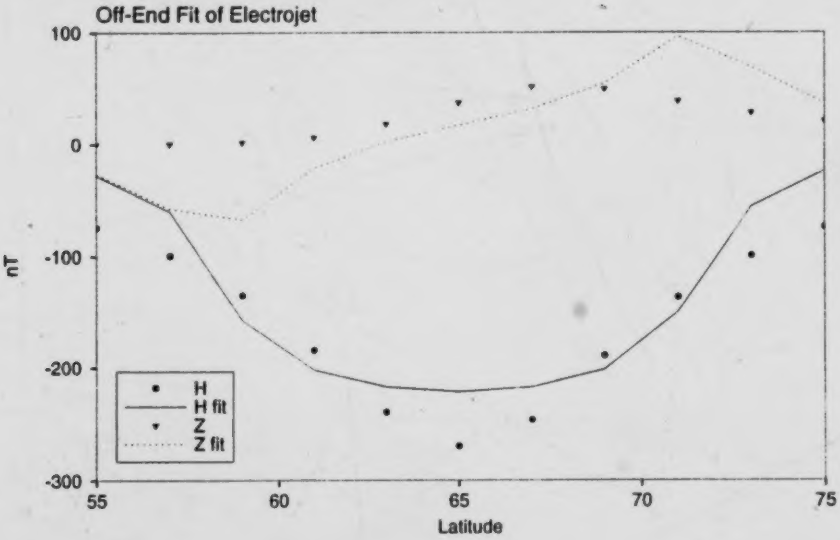
The determination of errors in a fitting procedure is difficult for several reasons. If the fit has fully converged to a minimum there is no linear error estimate in at least some parameters since the derivative of the error with respect to at least some parameters is zero at a minimum. If there is also difficulty in telling whether a global extremum has been reached, then there is no real way of knowing whether parameters which do not have zero derivative indicate a slope toward the true solution or simply indicate error. If there are a large number of parameters, as there are in geomagnetic modelling, these limitations prevent much quantitative assessment of error.

Another aspect of error analysis is the suitability of the forward model itself. If the forward model does not contain elements capable of representing the data, then the optimal model attainable through adjusting its parameters cannot be as good as that of a model which does contain those elements. An example would be an electrojet.

aligned east-west. With such an alignment there is no Y component in the central region. However, it is normal to find a north-south ionospheric current system associated with an electrojet, and such a system produces an eastward perturbation (although usually a small one compared to the X component due to the electrojet). The east-west aligned electrojet forward model would not be an adequate forward model to represent the Y component. Unsuitability of the forward model can also lead to ambiguity. If an electrojet forward model is used with data only from a latitude chain, both the presence of a north-south ionospheric current system and electrojet tilt could cause a Y component to arise. It is difficult to discern and separate the effects of these two sources of Y perturbation with a latitude chain. Thus a latitude chain would lead to ambiguous results. The situation could be made clearer with a two-dimensional distribution of stations (possibly two relatively nearby chains), in which case the electrojet tilt would manifest itself as a change with longitude of the latitude of maximal X perturbation (for example). In the absence of tilt, it would be inferred that north-south currents had produced a Y component. With the known properties of auroral currents, both sources of Y component (electrojet tilt and north-south current flow) would likely be present. With data from a two-dimensional region, the causative currents could be solved for, whereas latitude chain data alone would not permit this.

Yet another source of ambiguity of an electrojet model in a latitude profile is the production of negative X perturbations beyond the ends of a longitudinally limited westward electrojet (and positive X beyond the ends of an eastward electrojet). Such perturbations lead to a spurious equivalent current, giving a false impression that there is overhead ionospheric current. The presence of such perturbations was noted in the modelling of Kisabeth [1972] and they are largely the effects of field-aligned current. In Figure 5.6 such a modelling ambiguity is shown. The perturbation 10° east of the end of a 1 MA 180° electrojet lying between 63° and 67° latitude has been modelled by that from a centred electrojet (also of 180° longitudinal extent) whose latitudinal borders and current intensity were varied, in the manner described later in this chapter, to obtain an optimal fit. This represents an attempt to fit edge effects of an electrojet by a an overhead east-west current. In other terms, an equivalent current is being modelled through overhead current, where there is in fact none. It may be seen in the figure that the general trend of the X perturbation has been very well represented, but by an electrojet whose latitudinal borders (relatively well indicated by Z component extrema) were at 58.3° and 71.5°. The inferred current was 0.39 MA rather than 1 MA. Consideration of only the X component could mislead one into thinking that there is a rather wide (in north-south extent) electrojet overhead but the Z component behaviour has clearly not been well modelled. Rather than having two extrema as is the case for an east-west electrojet, the Z component has one maximum to the north and an ill-defined minimum or simple approach to zero to the south, a result consistent with that of Kisabeth [1972] for systems of smaller east-west extent. Not shown is the large Y component perturbation of the off-end current system, which resembles the forward model Z component. The forward model Y component would be zero in the case an electrojet centred on the latitude profile meridian. To distinguish a situation not well modelled by an electrojet, Z component perturbations must be considered. If these are small and Y component perturbations are large, the possibility of being off the end of an electrojet must be considered, especially if the X profile is wide. In global modelling as done in later chapters, the longitudes of current ends are allowed to vary (unlike in this example) and this problem should not arise. One should be aware of it in attempting to model perturbations in a meridian, however.

Figure 5.6 Perturbations off the end of a long electrojet, modelled by overhead current. The X component (circles) and Z component (downward triangles) 10° east of the end of a long electrojet have been modelled by a centred electrojet. The X fit is shown by a solid line and the Z fit by a dotted line.



f. Selection Criteria and Data Availability for Events

An initial criterion for event selection was availability of suitable magnetic data from widely distributed stations. In particular, since the method uses mid-latitude stations to constrain positions of field-aligned currents, the classic signatures of the substorm current wedge (bays of duration in the tens of minutes) were sought on midlatitude records. In the final selection, best magnetic data availability and that of other types of data for comparison were found to be from CDAW events. Analysis of the CDAW 9 events by other researchers is not complete and in that sense the analyses presented through automated forward modelling are new for those events. The CDAW 9A event has been most thoroughly modelled by other techniques and those modelling results are discussed for comparison. The CDAW 9B event has been recently modelled by AMIE [B. Emery, private communication, 1995] and global physical parameters are discussed as a complementary aspect of that modelling approach.

More recently, the GEM study of events from early November 1993 has supplied a suitable data set. The generally increasing number of magnetic observatories and the recent ease of interchange of data mean that near real-time inversion of magnetic data could be undertaken with Automated Forward Modelling on a global scale. On a regional scale it should be possible to actually implement real-time modelling. While political and economic considerations which cause gaps to arise in the station distribution (for example the collapse of the Soviet Union) can have adverse affects on any type of modelling effort. there is a continuing trend toward better communication, as exemplified by the rise of the World Wide Web (WWW). This suggests that near-real-time inversion could be done from data sets which come from all parts of the globe. The limit on such a monitoring program is the availability of suitable inversion routines, and this thesis presents one which could be suitable for the task. However, this is demonstrated here using only older data sets.

Data was initially available through specialized CDAW accounts at the NASA Goddard Space Flight Center, accessed by the SPAN dedicated network. Later, CDAW access was through the Internet. The CDAW accounts were eventually superseded by a CD-ROM, although still available for some data sets which were not included on it. Other data were obtained through direct contact with researchers who are listed in the Acknowledgments, and in particular through Barbara Emery at NOAA in Boulder whose gathered data sets were kindly shared. The individual data sets included those of the CANOPUS network. In the most recent months, geomagnetic and related data sets have been most readily accessed through the World Wide Web, which includes search tools which sometimes prove helpful in finding data sources.

g. Optimized Fitting of Magnetic Data

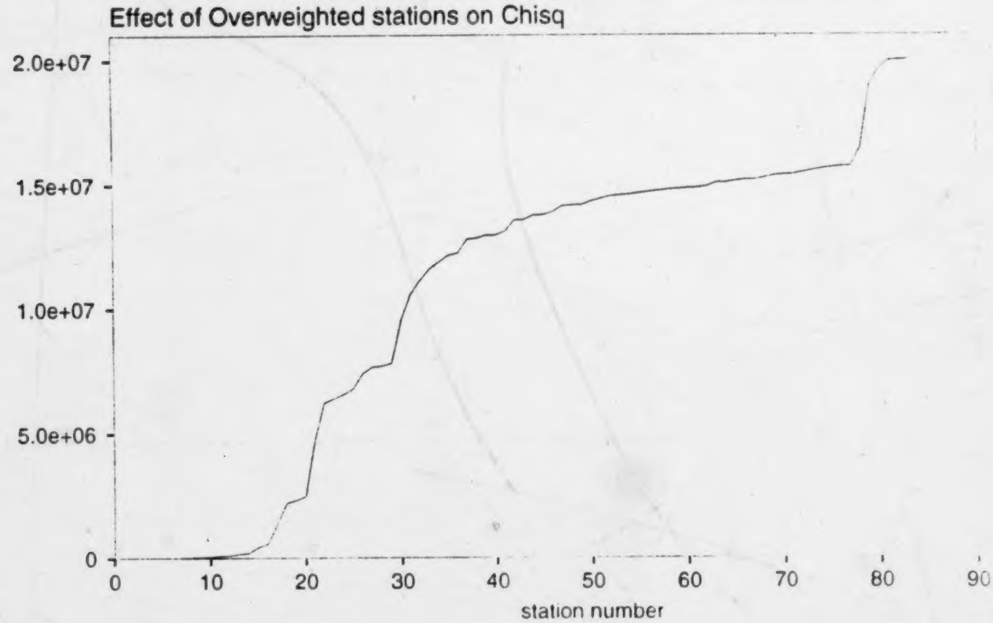
In the most general sense all selected events have in common activity in all of the auroral zone current systems, some of it intense. Perturbations are in each case visible at available mid-latitude stations. Depending on the rough position of the substorm current wedge as evidenced largely by D-component bays at mid-latitude and H-component bays in the auroral zone, stations were selected to give good coverage near the wedge. Distant stations are retained to act as constraints on the current system but need not be as numerous as in the area of rapid spatial change near the wedge. In practice about 30 stations were available for CDAW-based runs; for some events supplementary data were used and the number of stations could be increased to 80 or more. Regional studies can also be done based on 10 or less stations, and have proved particularly valuable if these stations form a latitude chain. The method is completely flexible in terms of number of stations, but judgment must be used to not allow the number of free parameters to become too great in relation to the amount of data available.

The automated forward modelling routine may be used on any subset of magnetic observations for which a reasonable forward model may be proposed. This can in principle include near-Earth satellite data or indeed multipoint observations with the aim of ascertaining the global configuration of the magnetosphere, but severe limitations in the amount of data available prevent such use. In practice, for several of the events treated below only a global ground-based data set is available with sufficient stations that the forward model may be used to explore details of the current distribution. In others a more general picture must suffice due to insufficient data. In this study, automated forward modelling of satellite perturbation data has not been undertaken and the few passes available are analysed roughly by inspection. In cases where enough data are available, the method may be applied globally and regionally and with models which are fully three-dimensional or essentially two or even one-dimensional. In particular, latitude profiles have formed the basis of many earlier uses of forward modelling [Kisabeth, 1972], and may be readily treated by the new automated technique. A polar-orbiting satellite pass has many similarities to a latitude profile but responds to an associated, but different, current system from that seen on the ground. Both are essentially one-dimensional situations and care is needed in comparing to the three-dimensional modelling. In both satellite and ground-based cases latitudes of current system boundaries are determined and latitude forms the important single independent dimension. In principle tilt of the current system could be determined and this would involve the orthogonal dimension. A forward model could even be proposed which incorporated field-aligned current and was essentially three-dimensional. Obviously, such an approach would have to be used with care not to overinterpret the situation. The effects of field-aligned current can resemble those of tilt of a current system and the ambiguity would have to be resolved by other means. One could also be led to interpret changes in latitudes of current system borders as indicated by one-dimensional modelling in a direct fashion where this is not appropriate. It is impossible to distinguish, in such a case, between actual latitudinal motion of the current systems, and possible east-west motion of a tilted current system. Once again, other means must be called in to resolve the ambiguity.

Having seen that considerable ambiguity can accompany one-dimensional modelling using latitude profiles, it may be useful to consider its advantages. Actual current flow through a meridian is readily established by such modelling if the latitudinal coverage is sufficient. In terms of automated modelling one can start with a poor fit to the data (as an initial condition, possibly derived by some very rough procedure which is readily automated) and arrive at a good fit. This is largely due to the reduced dimension of the modelling space. The three-dimensional approach has a very complex topology in a fit parameter space having typically about 30 dimensions. Initial conditions too far from the "actual solution" may not arrive at a globally optimum situation. Having reliable local conditions is a very effective way of constructing useful initial conditions which are derived from the data. In this way one can use initial one dimensional modelling along a meridian to form initial conditions for global modelling.

The use of weighted χ^2 as a fitting criterion implies that weights must be chosen. The absolute value of the weights can easily be seen to be of little importance as long as machine precision limits are not compromised by their choice. The relative values of weights determines to what extent goodness of fit at various stations creates a depression in parameter space. The programme attempts to find minima and will 'steer into' such depressions. Since auroral zone stations have larger perturbations (i.e. signal) but also more 'noise' in the sense that deviations from an ideal current system (as presented by the forward model) are nearby and therefore 'seen', they are weighted less than are mid-latitude stations which have smaller perturbations but also do not respond significantly to small irregular auroral zone variations. In practice the mid-latitude signal is about one-tenth of that typical in the auroral zone. To obtain about the same relative errors in the weighted χ^2, they are weighted about three times as much as are auroral zone stations. Sub-auroral zone stations received a weight about two times that of auroral zone stations. Depending on their strength and desire to attempt to model them, polar cap signals were weighted either the same or less than auroral zone stations. Figure 5.7 shows a practical way of determining that station weighting has been done properly by illustrating a case where it was not well done. The behaviour of χ^2 with station number (out of 83) is shown from an actual modelling run. It may be seen that it shows step-like behaviour with rather less variation in between the steps. This suggests that the stations at the steps are contributing too much to χ^2 and that those elsewhere may not be contributing enough. Such a diagram can be used to guide the choice of weights. It is not entirely clear that the weights would need to be changed in even such a case, however. The χ^2 will depend on deviations of the model fields from the observed fields, and these latter can vary enormously from station to station. In an early stage of the fitting process, a pattern such as that shown might even be expected and desired. Toward the end of the process, the χ^2 pattern should be smoother. Also, these considerations will affect the efficiency of the fitting computations, but in principle a fit can be reached despite poor weighting.

Figure 5.7 Behaviour of χ^2 as a function of station number.



To demonstrate operation of the fitting algorithm, a test with a simple current system is first described. The fitting parameter λ will be shown explicitly as it is changed by the algorithm according to the steps described in Section 5.c. The simple system consists of downward current along a meridian, ionospheric flow at constant latitude in a belt of finite latitudinal width, and upward current flow along a second meridian, as illustrated in Figure 5.1. Due to the simple nature of the system, parametrization in terms of longitudes, central latitudes, and latitudinal width of the current systems was used. In most other work presented here, the parameters (as presented in Table 5.1) are instead the geometric corners of the quadrilateral region delimiting current flow in the ionosphere. Table 5.3 shows the parameters of the test system and those of the initial guess. This current system is rather short in longitudinal extent, only 20°, and is of 5° north-south extent in the ionosphere, centred at 70°. This is very similar to current systems studied by Kisabeth [1972] and illustrations are given there of the perturbations from such a system. The guess deviates by several degrees in all parameters except one of the longitudes of upward current. The guess has about one sixth the total current of the target solution.

Table 5.3 Parameters of test (target) system and initial guess

	Long. Down 1	Long. Down 2	Latitude Down	Width Down	Long. Up 1	Long. Up 2	Latitude Up	Width Up	Current (MA)
Target	40	40	70	5	20	20	70	5	0.100
Guess	42	50	65.2	8	25.2	20	66	11	0.016

Figure 5.8 Chi-squared on linear scale throughout test convergence

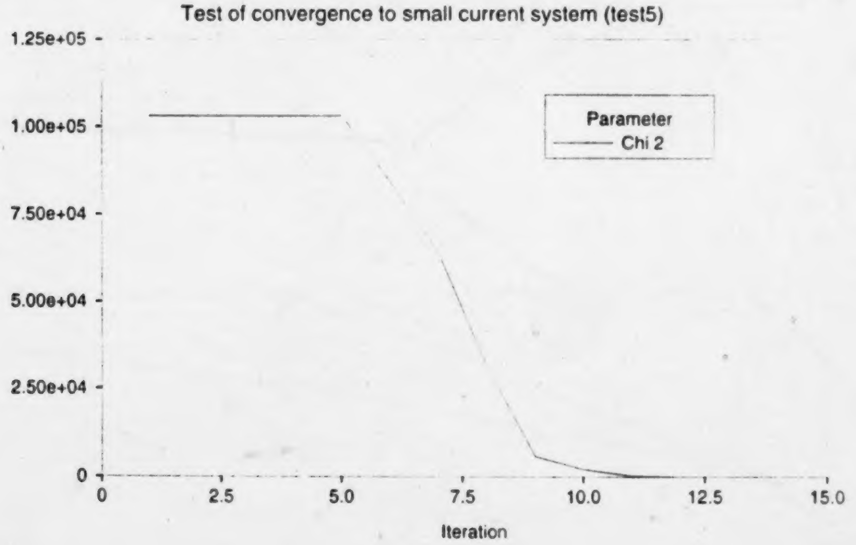
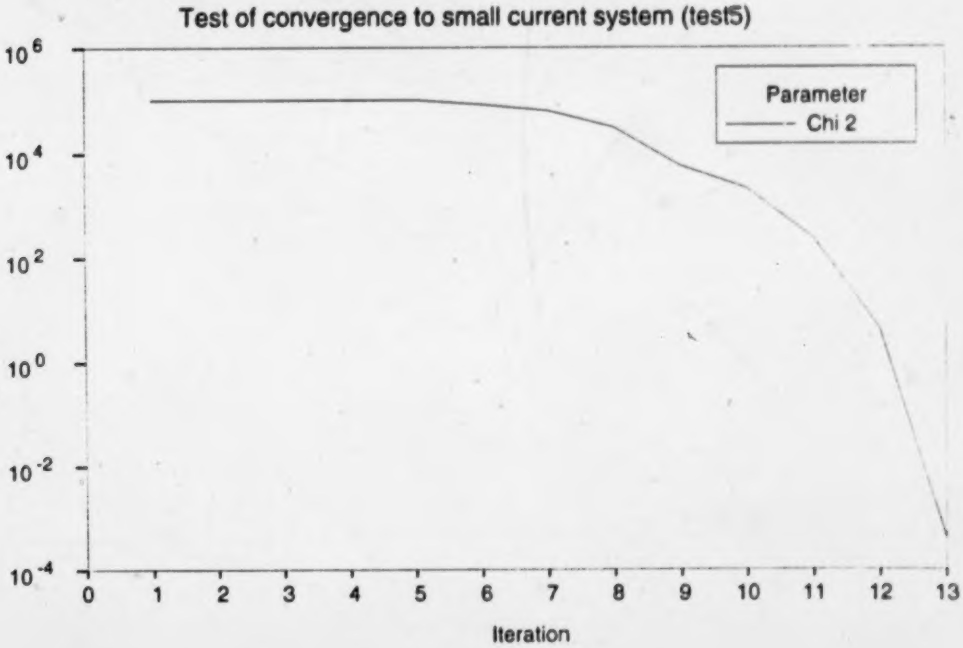


Figure 5.8 shows the values of χ^2 throughout this short run. Stations were selected on a grid covering the current system with one degree spacing in latitude, two degree spacing in longitude, and equal but arbitrary weighting was used at each station. χ^2 is thus an arbitrary number and its behaviour is more important than its actual value. It is seen to have values which initially remain as high as the starting value and after 5 iterations begin to decrease, going down by a factor of 10 after 4 more steps and then becoming too small to easily see on this figure.

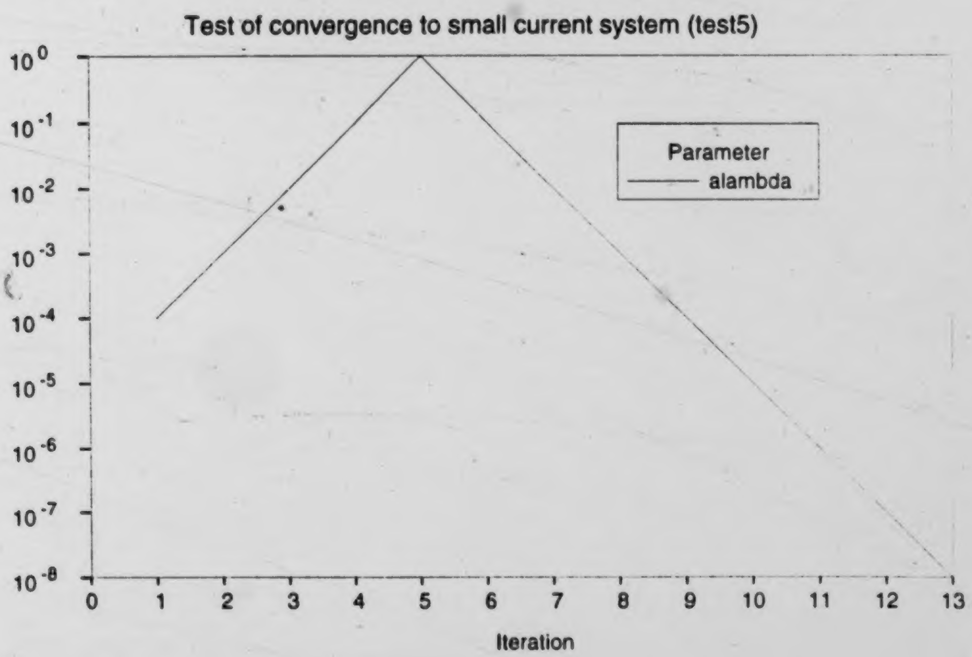
Figure 5.9 better illustrates the behaviour of χ^2 late in the convergence of this test case. It may be noted that the convergence, once it actually starts after step five, is better than exponential (which would be a straight line in this logarithmic graph). Also, χ^2 goes to extremely low values, reflecting the fact that there is no noise in this test run and that the forward model's parameters are exactly those of the 'data' and thus can be optimized almost arbitrarily well.

Figure 5.9 Logarithmic graph of χ^2 in test convergence



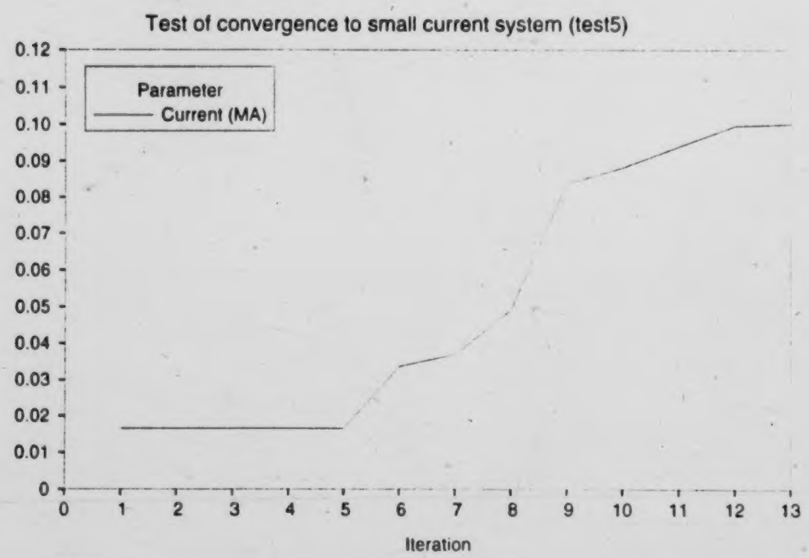
To better understand the mechanics of algorithm operation, the behaviour of the scale parameter λ must be examined, and is shown in Figure 5.10. As mentioned in Section 5 b, the authors of *Numerical Recipes* [Press *et al.*, 1992] seem to take a rather optimistic view of how close one is starting to a minimum, and recommend using a rather low value of λ initially so as to exploit the curvature (quadratic) aspects near one. In this case, initial parameter values are in fact rather far from those of the target, as is in part evidenced by the huge decrease in values of χ^2 through the run. At the initial stages, one is far from a minimum and the quadratic terms, brought into play by an inappropriate choice of the scale parameter, prevent moving toward the minimum, rather than helping. This is not directly seen in the graphs but may be easily inferred by recalling that if the proposed new parameter vector results in a higher χ^2 value, then it is rejected, and λ scaled up by a significant factor. The retention of a constant χ^2 for the first four steps as seen in Figure 5.9 or 5.10 simply means that the proposed new parameters, derived from solution of the **A**' matrix as detailed in Section 5.b, were rejected, and the old ones, with the accompanying value of χ^2, retained as the so-far-best guess. It may be easily deduced by inspection of Figure 5.10 that the 'significant' factor by which λ was scaled up was 10. Once λ has been adjusted to a value of one, at step 5, the χ^2 begins its steady decline, and λ is also adjusted downward, appropriately now as the new guesses are increasingly close to the minimum. Once 'over the hump' of choosing the right scale factor, the solution never slows down. The algorithm can rescale λ upwards once more if progress is not sufficient in some less ideal case, in hopes that such a re-emphasis on linear terms helps out.

Figure 5.10 Value of scale parameter λ during convergence of test case



The steady decrease of χ^2 might lead one to expect that the physical parameters would all show similar steady improvement toward their final target values. In the case of the current, perhaps because it is a linear parameter in the system, this is indeed the case, as is shown in Figure 5.11.

Figure 5.11 Convergence of total current in system from poor to good value



This parameter, initially far too small, is essentially doubled at the first effective step, and after that increases steadily, especially from step 8 to step 9 after (as will be seen) the other important system parameters of current latitudinal width and latitude have been brought near their target values. These nonlinear parameters vary in a more complex way during the optimization, as may be seen in Figure 5.12, which shows the latitudinal width of the system. Here also, nothing happens until the first effective step (step 5) and then changes occur in what we might initially consider to be the *wrong* direction, with the current widths, already too large in the guesses, made even larger. In a multidimensional nonlinear system it is sometimes difficult to visualize what is happening, which is a large part of the reason that we need algorithms like the one under discussion to help us navigate, but here we can usefully speculate. The initial latitudes (see Figure 5.13) were initially too low, and in fact the initially guessed system hardly even overlaps the target system. These are brought in the correct (northward) direction at the first step, the widths were increased, and the current was increased. The net result is that on average less magnetic perturbation was felt south of the target system where there was too much from the initial guess, more in the vicinity of the target system where there was initially too little, and the chi-squared decreased although some individual parameters might have become 'worse'. It is left to later steps (better due to being nearer the minimum) to correct this.

Figure 5.12 Uneven convergence of system width parameters toward target values

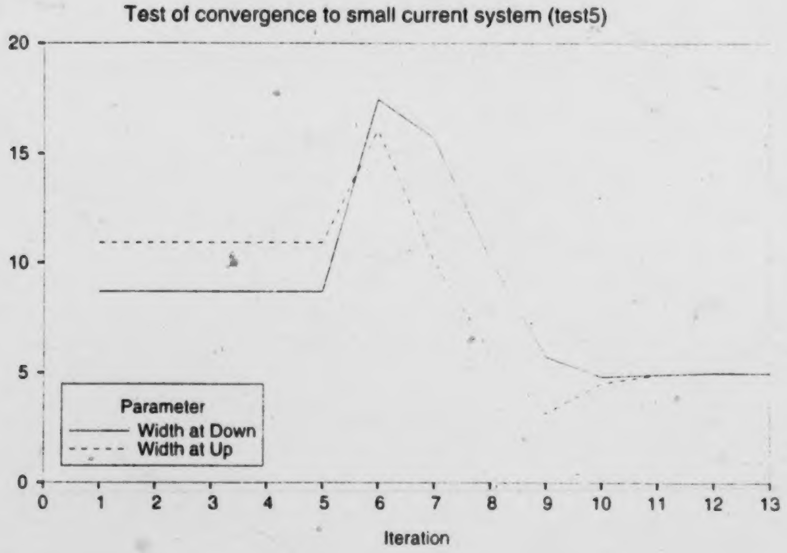
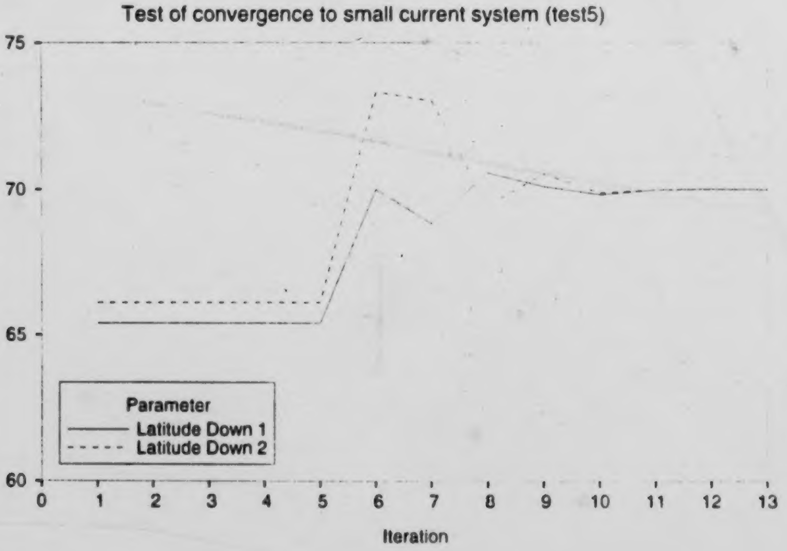
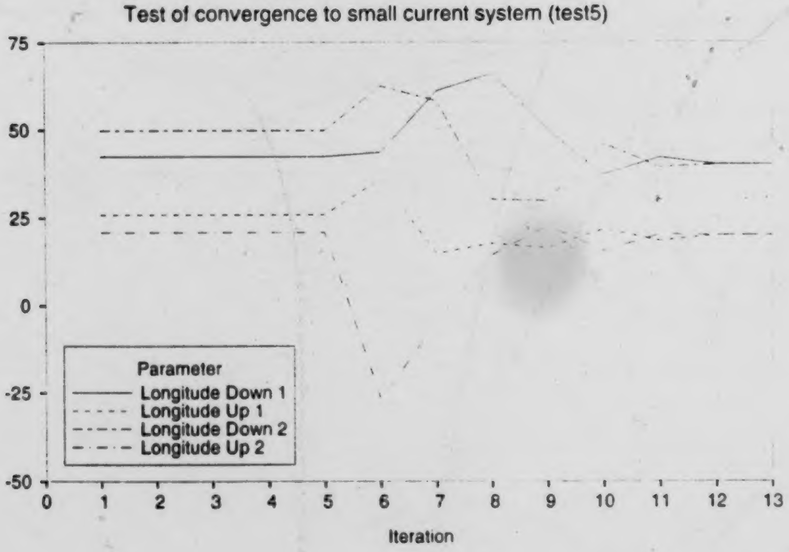


Figure 5.13 Variation of latitudes of centre points of up and down current sheets during convergence



The latitudes of the midpoints of downward current sheets at either end of the overall system are shown in Figure 5.13. Some overshoot occurs, particularly at the 'up-current' end. Although the target system is symmetric, the initial guess is not, and the nonlinear behaviour seems to have magnified that asymmetry at intermediate steps. By step 8 the latitudes seem to have reached values close to the target values. At the following step, both the widths and the currents make large changes to near their final values. It seems, not unreasonably, that the very important parameter of latitude had to have about the right value before the current and width can settle in on their best values. The various longitudes in Figure 5.14 reinforce the idea that overshoot can occur and that at intermediate steps certain individual parameters can get 'worse' before getting better. The astute reader will have remarked that, in Table 5.3, the initial parameter for 'Up Longitude 2' was in fact set to the known value of the target solution. It may be noted in Figure 5.8 that this 'fix' seems quite irrelevant: this value deviates widely from the initial (i.e. correct) value at steps 6 and 7, returning and staying near the target value at step 8. One's concepts of how to converge on a good solution must be relaxed considerably. This is one of two related problems faced in attempting to solve this sort of optimization manually [Connors *et al.*, 1991]. Changing individual parameters in the solution space can be quite misleading in itself, and the calculation of gradients in the space to guide such a choice (either explicitly or by 'inspection') is very tedious. Letting a proven algorithm do the work, with of course some inspection, is much more productive.

Figure 5.14 Variation of longitudes of extremal points of current sheets during iterations



The above discussion used one current system and illustrated changes in parameters during a solution, including the changes in the χ^2 and scale parameters which guide the behaviour of the algorithm. Due to this simplicity, once effective convergence started, it continued until convergence to the target resulted. It was stated that the algorithm can adapt to cases in which this does not happen: in effect that a certain amount of navigation around 'corners' in the solution space can be done. An example of this and of a further algorithmic step to double-check the optimality of a solution is now given. From an actual inversion of the data discussed in Chapter 6, Figure 5.15 shows a parameter which is used to track progress of a run and in practice decides when a run is finished. This parameter is a count of convergence failures in the run up to that point. With real data, the forward model will never be perfect, as it was in the simplified example above, and χ^2 will not become arbitrarily small. Rather, the minimum of χ^2 will have some finite value and once at that value, if it corresponds in fact to a global minimum, further reduction is not possible. Thus, further attempts at convergence will fail. For this reason it is useful to keep track of the number of failures to converge, since that may indicate that it is not productive to continue to attempt to optimize. It may also occur that there is never any improvement over the initial guess, and in that case the whole initial guess may need modification. Once the number of failures to converge exceeds five, the routine is once (only) restarted with a much larger ~~value~~ of scale parameter to ensure that one is not trapped in a local minimum, embedded within a larger scale parameter space slope which leads to a global minimum. After this check, and possible further convergence resulting from it, the result is accepted as a potential solution and calculations stop

Figure 5.15 Count of convergence failures in a full data run of the optimization. Decimal part of iteration number on ordinate is the number of iterations done. Abscissa indicates number of failures to converge since initializing the scale length

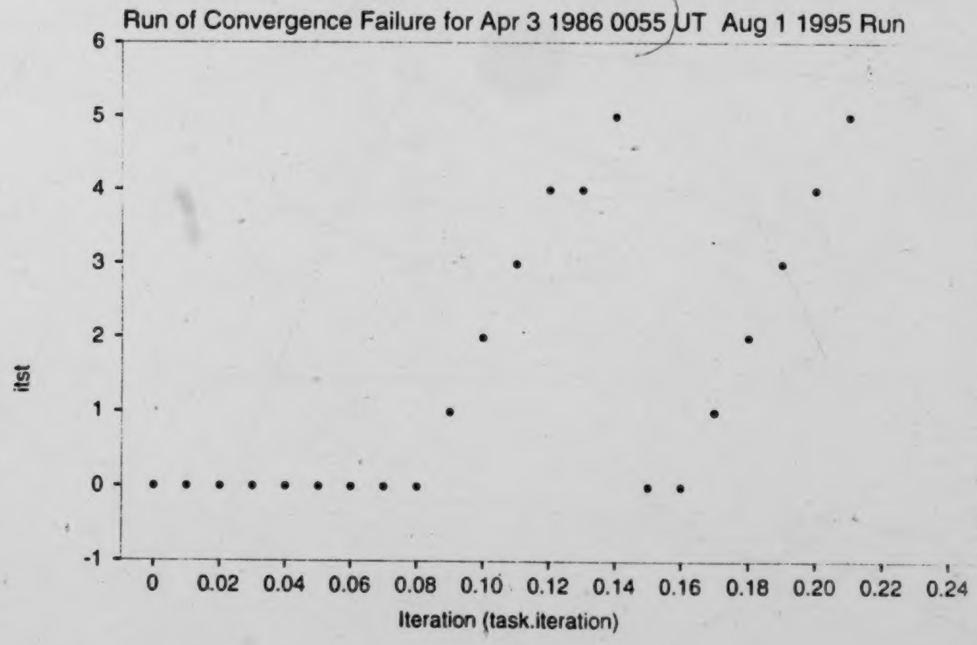
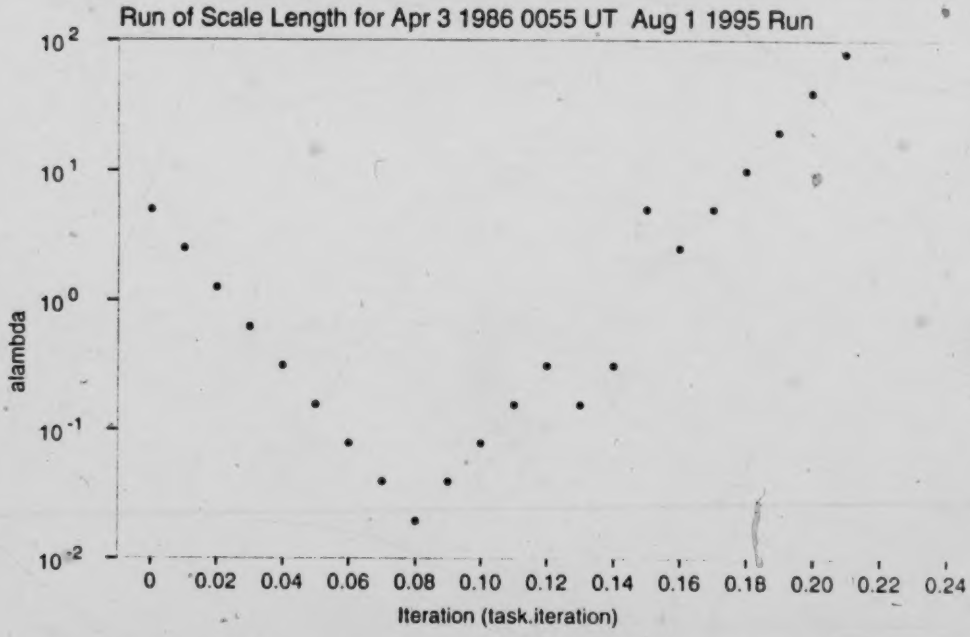


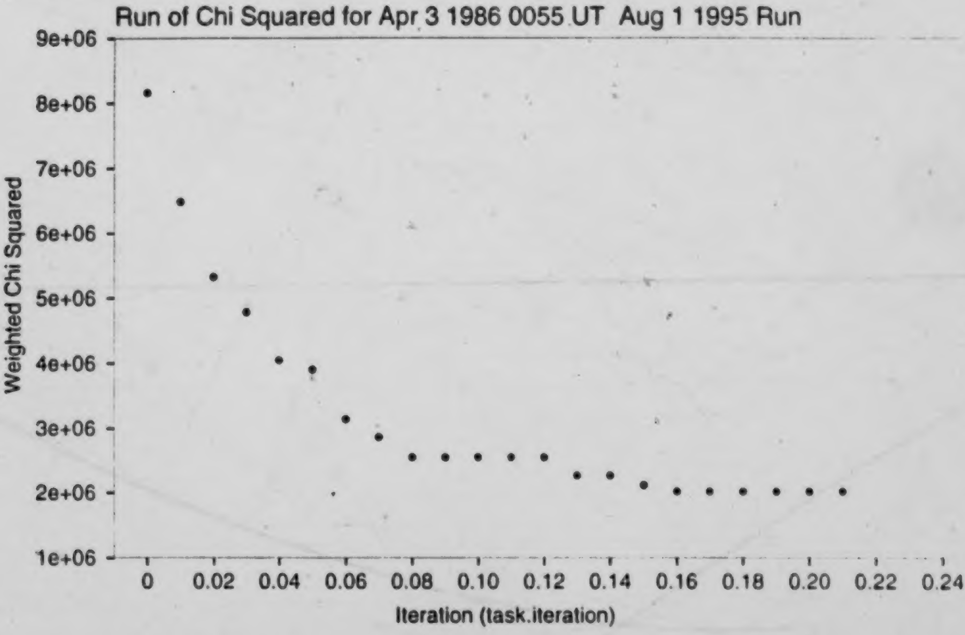
Figure 5.16 Scale length parameter λ in an actual program run. Ordinate labeled as in Figure 5.15. λ decreases each time convergence succeeds, increases on failure



The convergence failure count is best interpreted in conjunction with the behaviour of scale length shown in Figure 5.16. The initial lack of convergence failure means that useful convergences were being obtained; that is, that the new parameter values were closer to the minimum than the guesses preceding them and thus were accepted. The scale length in such a case is to be decreased and this is seen to be what took place. Experience has shown that in these problems the initial guess is not very good and that a large value of λ (in fact 10) is appropriate. The best value for the 'large factor' by which λ must be scaled proves to be 3. Improvements continued to be made, and no failures recorded, until step 8. At this point a failure to improve χ^2 occurred, and λ was scaled up, to a larger value in case convergence failure occurred due to being on a slope rather than actually near a minimum. This did not improve the situation except at step 12, and after 5 such attempts it was concluded that one is in fact near the optimal solution. In case the scale is still not large enough to detect being in a local minimum on the edge of a large slope in parameter space, a new set of iterations is attempted with the much larger value of λ that was initially used with the first guess. This does produce a small improvement, so the failure count in this second run through remained 0 and λ was decreased as per the recipe. However, no further improvement occurred even when λ, after one failure, was once more increased in value.

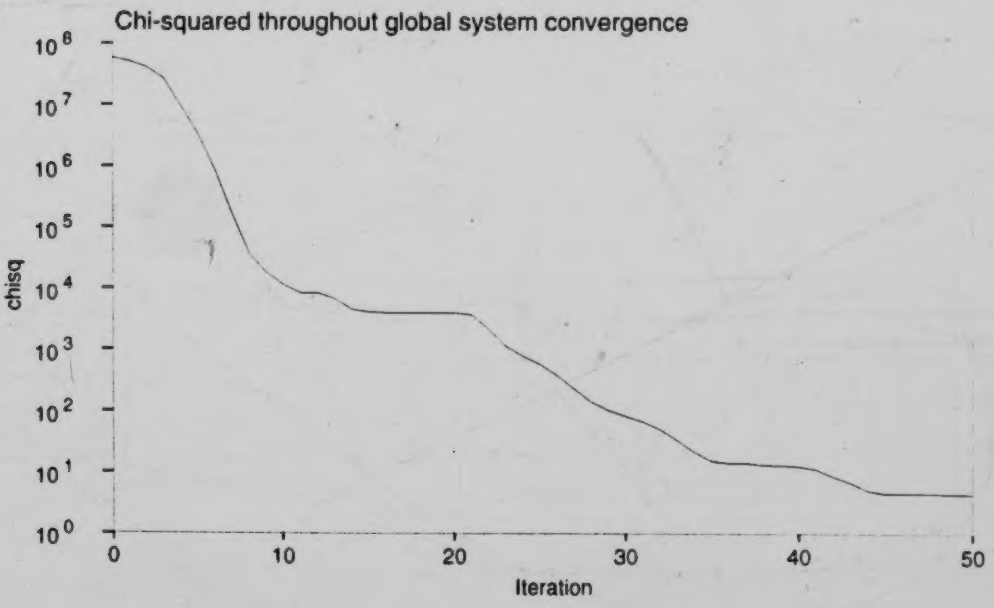
After a string of five failures to reduce χ^2 with yet larger values of λ, it may be safely concluded that the solution had been attained since the slope in χ^2 space on many scales was zero. Figure 5.17 shows the initial large decline in χ^2 and its subsequent flattening when near the optimal set of parameters. The set of parameters at the end of such a run may with some confidence be accepted as being a best fit of the forward model in use to the data.

Figure 5.17 Behaviour of χ^2 through a real optimization run. Ordinate labeled as in Figure 5.15.



As has been discussed above, generally nine parameters are needed to characterize each toroidal current system, and generally three systems are used to represent global perturbations. Thus, at least 27 parameters are needed to represent the data. At least one other ('D_{st}-like') parameter must be added to represent low latitude perturbations, likely due to external current sources. There may also be free parameters associated with the poloidal current systems, although in most cases one would use colocation (see above) to reduce the number of parameters associated with these systems. By constraining the geometric parameters to be close to those of the toroidal systems, only the current in poloidal systems remains as a free parameter. This means that 31 parameters should suffice to represent the global current systems. In principle, with three components per station, about 10 stations should suffice (with essentially a one-to-one parameter remap from magnetic data space to current parameter space) to determine these 31 parameters. In practice, for global modelling, more than 30 stations must be used in order to have enough data to determine these parameters with some degree of accuracy in global modelling. Although each station provides 3 components of data, some stations are not well placed and contribute little information. In the most complete global modelling presented in this work, 80 or more stations are used. The good fits attained imply an economization by about a factor of 8 through use of current parameters as opposed to the raw data. On the other hand, the 31 (or more) dimensional current parameter space must be considered 'large-dimensioned' in several practically important fashions. Visualization of data and model parameters in this space is itself a challenge. Further, it is obviously difficult to be sure that any extremum found in the space is truly a global extremum. Even crude tests of the extrema, such as taking steps away from a purported minimum, are difficult given the large dimensionality of the space. For these reasons, it has been pointed out that one must carefully survey the steps taken to get to a minimum, and the overall behaviour of the programme while it does its optimization. Further, one must understand the operation of the Levenberg-Marquardt algorithm, as described above, to interpret the behaviour. It has been shown that one should not attach too much importance to the changes in individual parameters as the optimization proceeds, provided that both χ^2 decreases in a systematic fashion, and that the final solution produces perturbations which are comparable to the input data. As a final example/test in this section, a 'poor' initial guess is used with an idealized yet realistic forward model. This demonstrates that an initial guess can be far from the target and that the routine can still obtain convergence, in other words, it is a demonstration in principle that the method under discussion works well on the type of data available.

Figure 5.18 χ^2 throughout convergence of a realistic but 'poor' initial model



In this test, 27 parameters were used to simulate the Hall currents associated with an idealized auroral oval current system consisting of morning and evening sector electrojets and an active westward electrojet in the midnight sector. Neither a D_{st}-like parameter nor poloidal currents were used in this illustrative example. The surface magnetic fields were calculated at actual magnetic observatory locations using the target parameters. These 'observations' were then used as input to the Automated Forward Modelling routine. Since the main aim here is to demonstrate that a poor initial guess leads to convergence even with many model parameters (in other words that the AFM routine works in the ideal case), the run of χ^2 throughout convergence is shown as Figure 5.18. Much as in the single current case (see Figure 5.9), there is seen to be rapid convergence.

Figure 5.19 shows a global view of current systems put forward as a reasonable representation of the auroral oval currents, yet one which deviates considerably from the target system. The very large initial values of χ^2 seen in Figure 5.18 attest to the fact that the two systems are considerably different. Station weightings varied in the calculation of χ^2 much as in the real calculations detailed in the following chapters, with greater weight given to mid-latitude stations. As mentioned above, the scale parameter is of importance in attaining a solution: when this parameter was set initially to 0.001, no progress was made toward solution. Such a small value is suitable if one knows that the parameters of the initial forward model are very close to those of the solution [Press *et* al., 1992]. Here that is not the case and the curvature terms which are emphasized at small values of the scale parameter dominate when it is not appropriate for them to do so. Progress to solution was obtained by setting the initial scale parameter to 10. This is also the starting value for this parameter which was generally used in applying AFM to real data. The test case solution after success of AFM is presented as Figure 5.20. Currents solved for were roughly a factor of two changed from those in the initial model, and the latitudes are seen to be considerably higher than those of the initial guess (Figure 5.19). The active midnight sector has been found to be at much different longitude that has that of the initial guess, and the noon sector downward current configuration is also considerably different. The final solution's parameters differed only at the hundredth of a percent level from those of the target system, consistent with the very small χ^2 values attained. This constitutes a convincing demonstration that the AFM method works well in the mathematical sense: it is capable of inverting data supplied to it. Further, in this ideal case, the routine is seen to be efficient in terms of number of steps taken to do the inversion.

Figure 5.19 Initial current configuration for realistic global test model

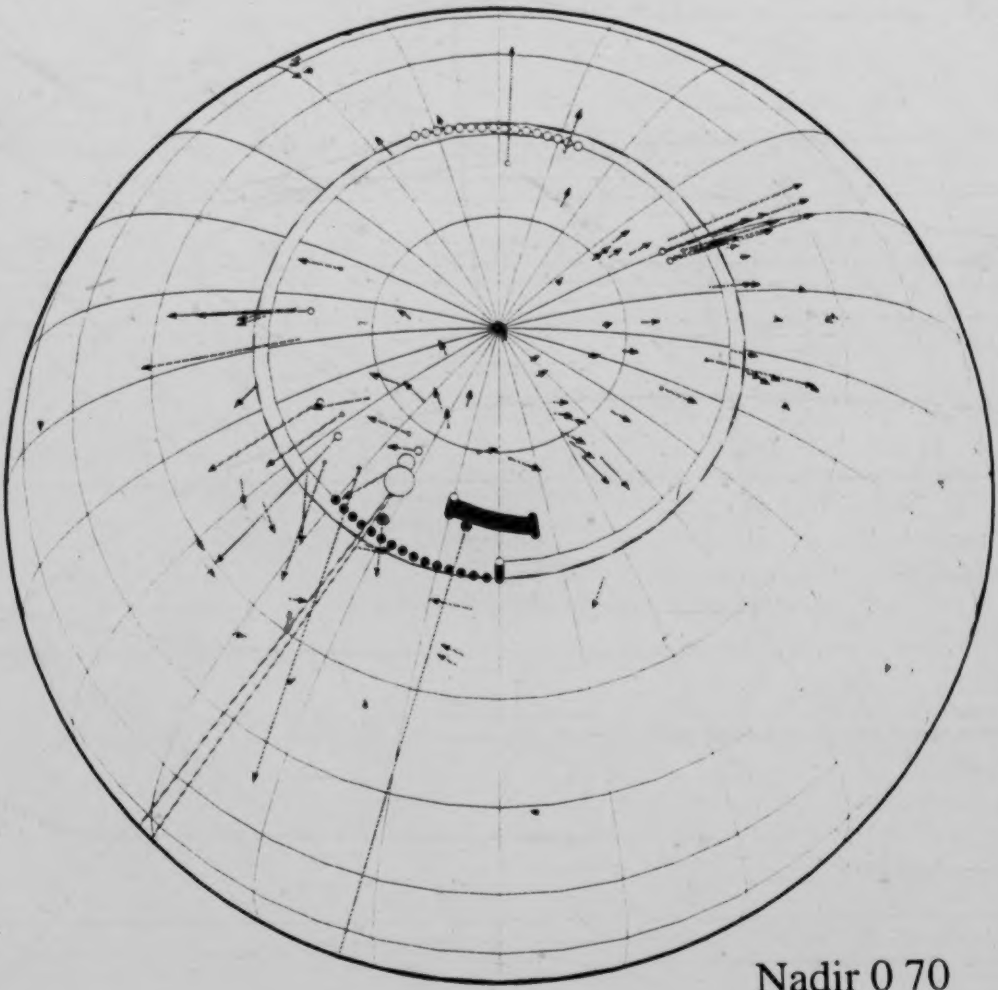
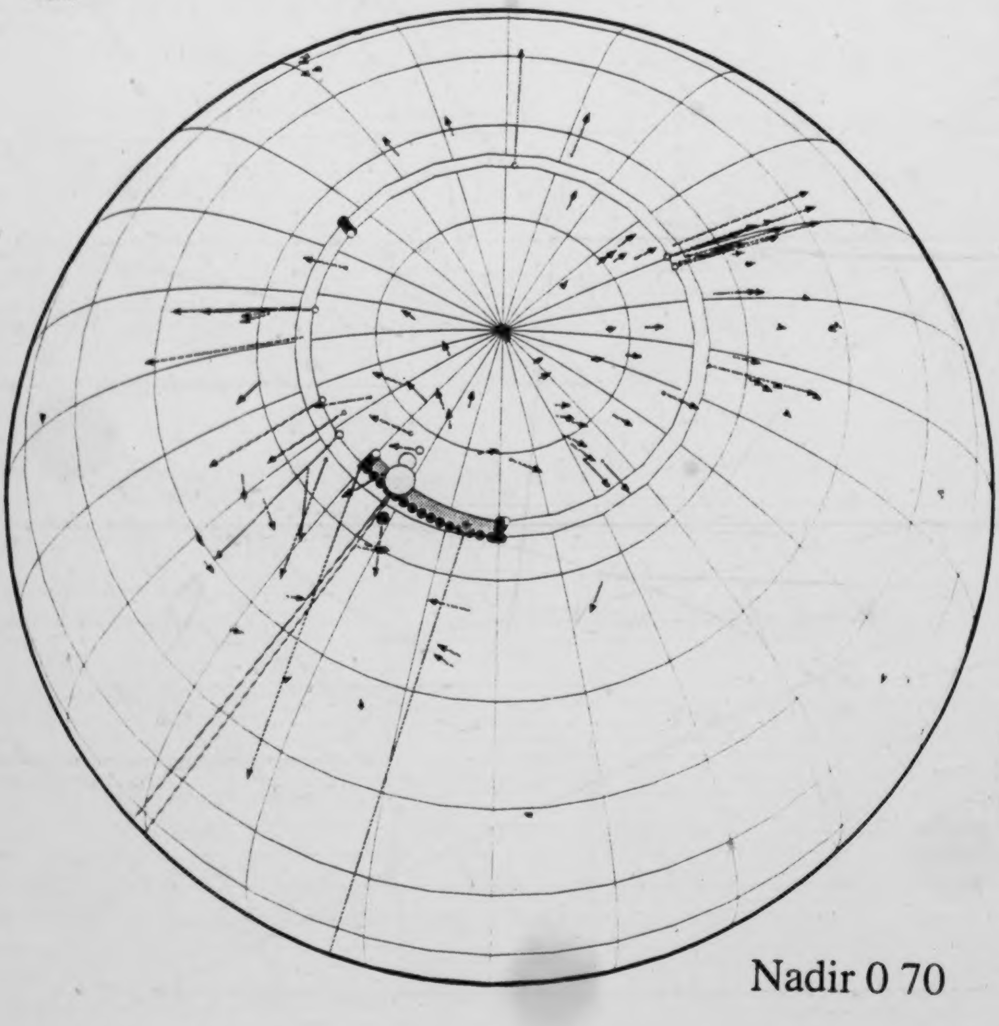


Figure 5.20 Final solution, showing perturbation magnetic field, for AFM global run.



Nadir 0 70

h. Comparison with Other Data about the Events

In several cases in what follows (Chapters 6, 7, and 8), images from a spacecraft (Viking) allow rough determination of particle precipitation regions. Comparison of these regions with those deduced from the magnetic data are particularly useful. In all cases, data from various other spacecraft are available and are used to discuss the larger context of the events, and to a limited extent implications for mapping of field lines. DMSP magnetic and particle data can be used to verify results due to the known colocation of the Hall electrojets which produce most of the ground signal and the Region 1/2 field-aligned currents which produce the perturbations dominant in near-Earth space. In certain cases it is of particular interest to determine whether or not a substorm onset occurred and the deciding factor then may be other signatures of onsets such as near-tail dipolarization or particle injections.

The various types of other data available were discussed in Chapter 3. The forward models presented here are computed using only magnetic field data. Other forms of data have not been incorporated directly into the modelling, but rather used to provide a broader context to discussion of results. In principle the AMIE method (see Section 4.b) allows incorporation of certain other data since the model which is part of that inversion scheme is more physically complete than that used here.

Chapter 6. CDAW 9B - April 2-3, 1986

This is the first of several chapters devoted to analysis of auroral zone activity using the Automated Forward Modelling techniques described in Chapter 5, in conjunction with other data. Other events studied, selected in accordance with criteria of Section 5.f, are discussed in separate chapters, and all are listed in the Index. This particular event is analyzed in more detail than any others considered.

The activity of interest spans the late UT hours of April 2 1986 and continues until 1000 UT on April 3. During that time several substorm expansive phase onsets occur. Such onsets begin at roughly 2000 UT April 2 although there was some less clearly defined activity prior to this time also. The activity is shown through Viking images and ground magnetograms but is not properly indicated by the AE index. Detailed analysis will be presented of Viking images, available from roughly 2030 to 2120 April 2 (orbit 220), 0100 to 0200 April 3 (orbit 221), and 0545 to 0625 April 3 (orbit 222). In the earlier orbit there is evidence of high-latitude substorm activity which is consistent with the magnetic observations from Greenland. In orbit 221 the development of upward field-aligned currents can be correlated between images and magnetic modelling. The period around 0100 has been considered by Skone [1994] with particular attention to changes at geostationary orbit at substorm onset (0055 UT). A large onset occurs at approximately 0220 but is not covered by Viking imaging. An onset is clearly visible in Viking imaging at 0615 but its development cannot be followed due to cessation of imaging shortly after onset. This event has been discussed with emphasis on recovery phase structures (Ω bands) by Opgenoorth *et al.* [1994], who identify two substorm expansive phase onsets (0100 and 0151 UT) based on magnetics, imagery, and Los Alamos National Laboratory geostationary spacecraft observations of drifting energetic particles observed shortly after these inferred onset times. It will be shown here that in fact many onsets occurred during the period in question. Besides the ones mentioned above, others occurred earlier and at high latitude, and an important morning sector onset at high latitude at about 0130 will be discussed in detail. Where it is not clear by context, or unless explicitly stated otherwise, times in the text which follows refer to UT on April 3. Also the two ISEE spacecraft are very close together during this event and if the term ISEE is used alone, it refers to both spacecraft and to near-identical signatures at both.

a. Chronological Summary

The actual Automated Forward Modelling global results presented here focus on the period 0030 to 0300 UT on April 3 but are put into perspective by examining data from outside this period of optimal magnetic data, using other sources of data and regional magnetic modelling. The following series of tables summarizes activity. Each table corresponds to a section of this chapter, but their presentation together here permits a global overview.

Table 6.1 Conditions preceding activity (Section 6.b)

Time (April 2 1986)	Activity	Observational Basis
1100-1300	high latitude onsets and polar cap precipitation	Viking, ground magnetics
1600-1800	onset-like activity near Dixon; high latitude electrojet in Greenland	ground magnetics
just before 1800	low level of magnetic activity and marginally southward IMF development of eastward electrojet	AE index, IMP-8 Greenland stations at 75° invariant latitude
2005-2045	sharp southward turning of IMF Increase in electrojet strength as seen in Greenland and Soroya	IMP-8 ground magnetics
2045-2100	abrupt change to northward IMF Activity at base of sun-aligned arc near Greenland Electrojet active in E. Greenland consistent with brightenings seen by Viking Abrupt H drops at Soroya with D peaks	IMP-8 Viking image ground magnetics, Viking ground magnetics
2100-2140	near-zero IMF B_z continuing auroral brightenings active electrojet	IMP-8 Viking ground magnetics
2140-2155	northward IMF of small magnitude, ca. 2 nT	IMP-8

Table 6.2 Start of active period (Section 6.c)

Time (April 2 1986)	Activity	Observational Basis
2155	definitive southward turning of IMF abrupt decrease in SCO X and Z component magnitudes	IMP-8 ground magnetics
2200-2220	abrupt onset at SCO upward FAC in eastern North America	X component to -175 nT Pi 2 at Syowa minor H bay at Isafjordur increase in D component at Saint John's
2240-2315	abrupt onset at SCO (2245) electrojet intensification on W coast of Greenland	X component to -125 nT Pi 2 train and noise at Syowa small electron injection at SC 1984-037A minor H bay at Isafjordur no change in AL or AU GHB X to -40 nT
2315-2340	activity decrease in Greenland steady level at Soroya little change in AU or AL	ground magnetics " " " "
2340-2400	complex SI/SSC likely associated with solar wind pressure pulse	H positive/negative signature at Soroya AU bump, AL starts decline bump in northward field at equator
Time (April 3 1986)		
before 0050	strengthening of morning sector westward electrojet undulating northern border of evening sector electrojet 'growth phase' decline in proton fluxes in morning and midnight equatorial plane -Y and -Z 18 R_E downtail	ground magnetics Viking Los Alamos synchronous satellites ISEE at GSM Y=-12
0045	electron dropout/injection	S019 synchronous satellite

Table 6.3 First major onset (Section 6.d)

Time (April 3 1986)	Activity	Observational Basis
0050-0120	formation of substorm current wedge at northern border of westward electrojet	-D bay over EISCAT and Europe; +D in North America
		major proton injections: dispersionless at S019; dispersed at S037 at 0100 UT
	extension of westward electrojet across Atlantic	D and Z components in NA
	westward stepwise expansion of the wedge	Viking
	evidence of expansion of Pedersen system to near GOES 5	GOES 5, GOES 6
	end of ISEE -10 nT B_Y, -5 nT B_Z perturbations at 0055	ISEE at GSM X=-18, Y=-12, Z=-2.4
	short-lived 'wedgelet' formed N of PDB at extreme position of wedge	PDB large, discrete Z/H perturbation bright spot over Hudson Bay (Viking)
0120	UV spectral changes when wedge dies out	Viking

Table 6.4 Northerly onset (Section 6.f)

Time (April 3 1986)	Activity	Observational Basis
0130-0150	substorm current wedge at northern border of expanded westward electrojet allows its further poleward expansion	onset at Bjørnøya +D over N Atlantic; -D over widespread area of Europe Viking imagery of far north substorm "eye" dipolarization in morning sector (AMPTE IRM, ISEE) electron/proton injection at S019 ISEE keV particles increase
	Ps 6 over EISCAT	ground magnetics
0150	onset in narrow latitudinal range near 72° in W. Greenland	FHB magnetogram pulsations/bay at Iqaluit brightening in Viking images

Table 6.5 Second major onset (Section 6.g)

Time (April 3 1986)	Activity	Observational Basis
0150-0220	slow intrusion of morning sector electrojet further into evening sector distant field more tail-like	PDB and E North American magnetics ISEE at X=-18
0215-0240	major substorm current wedge south of morning sector intruded electrojet poleward motion so that activity moves toward BLC strengthening/lengthening (or southward motion?) of evening sector electrojet westward motion of wedge	-D bays superposed on +D in E. NA; -D bays in W North America; onsets nearly simultaneously at FCC and PDB; electrons at S019 and S037; protons at S129 (evening sector); dropout of electron flux at S129; dipolarization at ISEE ground magnetics in midnight sector BRW magnetogram transition bay at GLL
0240-0250	continued northward motion of wedge and extension eastward to IQL	ground magnetics

Table 6.6 Recovery (Section 6.i)

Time (April 3 1986)	Activity	Observational Basis
0300	strengthened electrojets	ground magnetics (end of Emery data at this time)
0305	northward turning of IMF	IMP-8 (end of IMP-8 data)
0500	cessation of negative H bay at BJN	BJN analogue magnetogram

Table 6.7 Later activity (Section 6.j)

Time (April 3 1986)	Activity	Observational Basis
after 0300	general decrease of activity until ca. 0600	auroral indices and ground magnetic data
0615	substorm onset south of Gillam	Viking, Gillam magnetogram, particle injections/dropouts

Figure 6.1 IMP 8 solar wind data for April 2 1986. From top: total field, Z GSM component, and GSM Longitude of field vector. Data gap from about 5 to 12 UT.

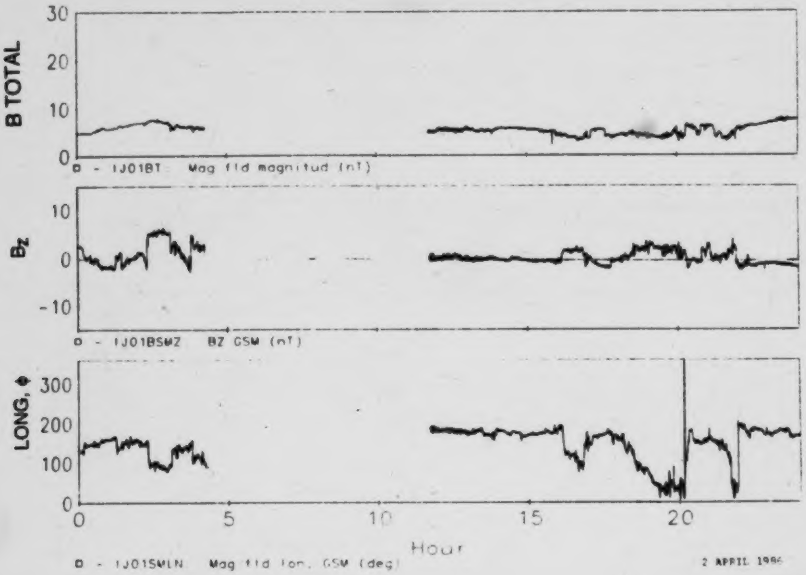
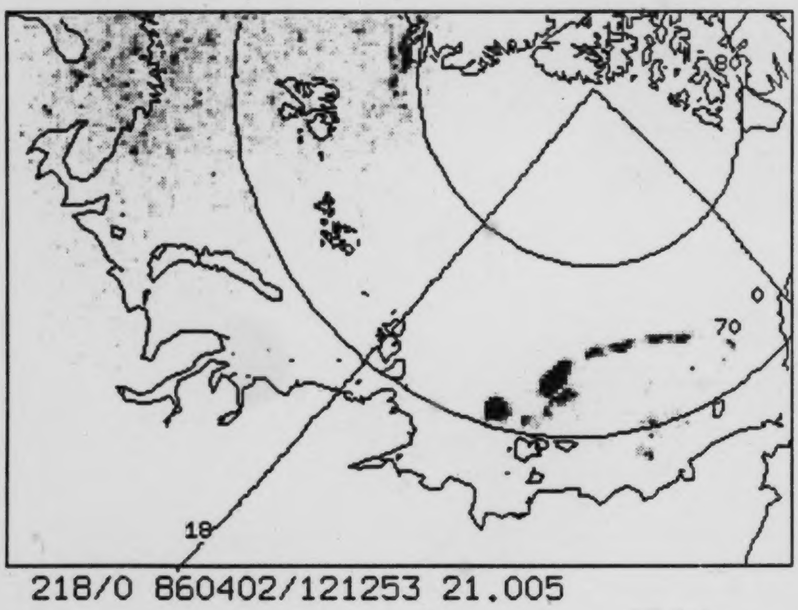


Figure 6.2 Viking view of expansive phase activity on April 2 1986.

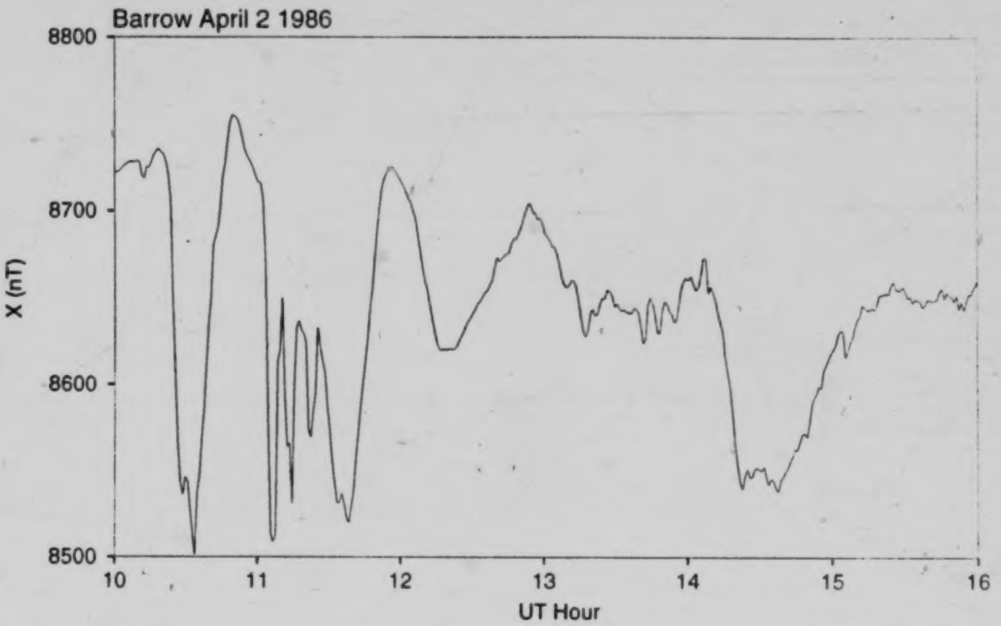

80
70
18
218/0 860402/121253 21.005

b. Conditions preceding activity

The period well before the interval of detailed analysis has numerous interesting aspects and is useful in assessing conditions leading up to the major substorm onsets listed above. IMP-8 solar wind data are available starting at shortly before 1200 UT April 2. Solar wind data for this early period are shown in Figure 6.1. Once available, they indicate a prolonged period of near-zero but predominantly slightly northward (and essentially radial) field until about 1630 when the field became more northward and returned to a more normal firehose angle (about which there were subsequently large fluctuations in direction). A period of unsteady field followed until a decisive southward turning at about 2200. Some expansive phase activity is suggested by the AE index before 1200. Inspection of Figure 6.2 shows that a further substorm expansive phase occurred shortly after 1200 when the IMF northward component was small or northward. Viking imaging commenced at 1159:59 UT and shows the development of a complete expansive phase. The Viking data are shown by Murphree *et al.* [1991] in false colour, and they discuss the phases of this isolated event and the fact that its AE signature was relatively small. This expansive phase is not adequately indicated by the AE index (refer to Figure 6.7) since it is at high latitude: indeed, inspection of Russian magnetograms (data not shown) shows that perturbations in that sector, containing two possible AE contributing stations, are minimal.

The magnetic signature of this expansive phase is best seen at Barrow (Figure 6.3) where a typical signature commences with a steep X component decrease at or just before 1200 UT. The Z component during this X component bay was about -50 nT, or about half the X perturbation. This would indicate that the associated electrojet was well north of the station. The magnetogram also indicates another onset, accompanied by strong pulsational activity, at about 1100 UT (and another at about 1020). The 1100 onset is shown in AE, including distinctive pulsations. While it is unclear, due to lack of IMF information, how these onsets related to the IMF, it appears that the 1200 UT onset took place while the IMF was northward and likely also commenced during northward IMF. Disregarding the possibly unusual IMF condition (see, however, the discussion by Henderson *et al.* [1996]), let us simply state for the moment that this event occurred at high latitude and involved a current larger than would be estimated from the X or H component alone at any station, or by extension by AE. Further inspection of the Barrow magnetogram shows a second onset-like feature at 1400 UT with the bay continuing beyond 1500. If this is a true onset, the IMF had been northward or near-zero for at least two hours prior to it. High-latitude activity on the contracted auroral oval appears not to necessarily be associated with southward IMF.

Figure 6.3 X component at Barrow from 10 to 16 UT April 2 1986



c. Start of active period

The period just preceding the nominal CDAW interval (beginning at 1800 April 2) features low amplitude southward turnings of the IMF (see Figure 6.1). During this early period disturbances indicating a westward electrojet north of the station Dixon are seen, and an eastward electrojet is observed at latitudes as high as 80° in Greenland (data not shown). As shown in Figure 6.4, UV auroral emissions, which usually generally indicate the location of the electrojet, appear to rise steeply in latitude in going from the midnight to the evening sector. The evening sector auroral oval is at high latitudes in Greenland, toward the left in the images, at up to nearly 80°. The midnight MLT meridian is marked on two of the images. It is seen that the auroral oval near midnight is at considerably lower latitudes than in the evening sector, reaching nearly 65°.

Disturbed conditions during the early CDAW period are clearly associated with southward turning of the interplanetary magnetic field (IMF), shown in Figures 6.5 and 6.6 from measurements at IMP-8, located tailward of Earth in the evening sector well outside the magnetopause. The spacecraft GSM coordinates change from (-17.6,32.7,-7.8) to (-12.2,34.3,-10.1) from 1800 April 2 to 0300 April 3 (after which there is a large data gap). A brief excursion southward accompanies a major change in the GSM X component (along the Sun-Earth line) at 2010 UT April 2. The Z component recovers to positive or near-zero values after approximately one half hour, while the X component remains negative until about 2130. Its abrupt reversal at 2150 is accompanied by a southward turning and from this time until nearly 0300 UT April 3, B_Z remains negative. Y component changes accompany the two major B_X changes but are of smaller magnitude: B_Y is positive during much of the period considered but near zero in the 2150 to 2330 time frame. IMP-8 data are not available after approximately 0305 on April 3: there appears to be an abrupt northward turning (accompanied by a decrease in GSM Y component but not much change in the X component) immediately before this time. Data available once more at about 1000 UT and onward show a near-zero Z component becoming strongly positive for most of the latter part of the UT day.

Figure 6.4 Viking images from orbit 219, April 2 1986. EDFL coordinates are shown and images have been contrast-enhanced. Image at lower right was taken with lower camera gain and shows only near-midnight forms.

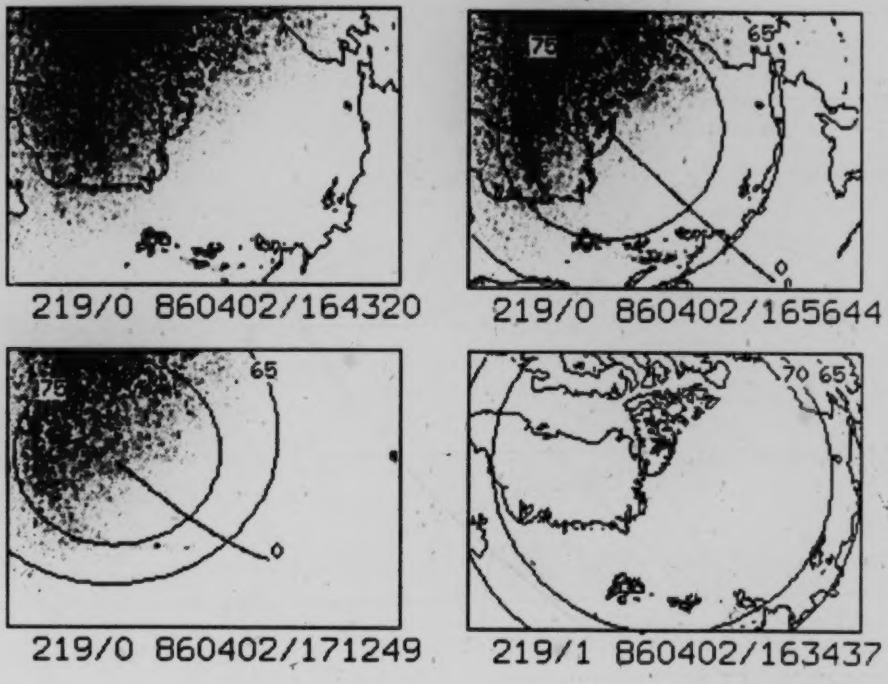


Figure 6.5 IMP-8 magnetic field on April 2-3 1986. GSM X (solid) and Y (dashed) components are shown.

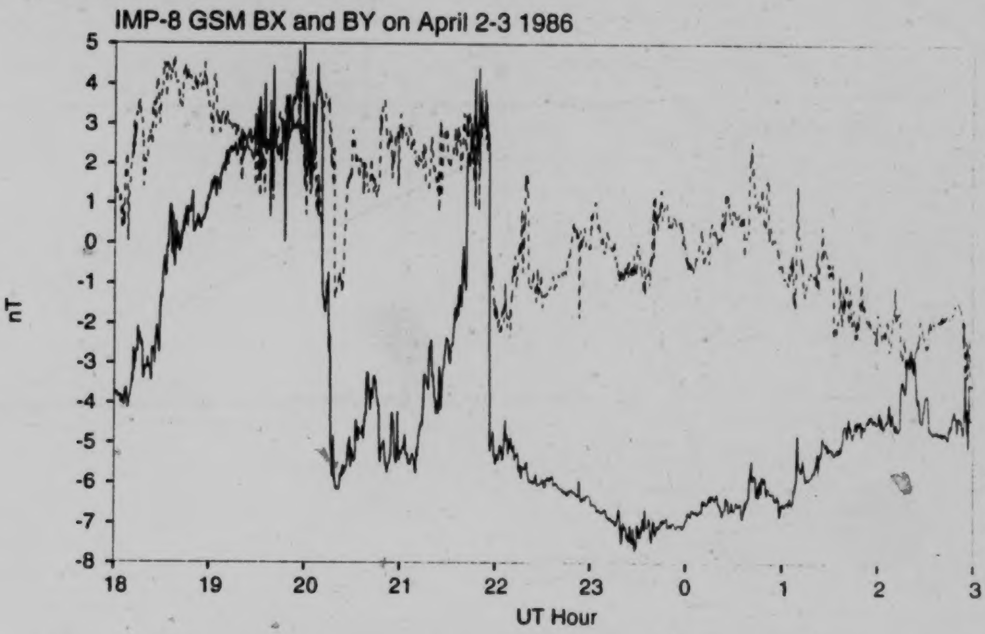
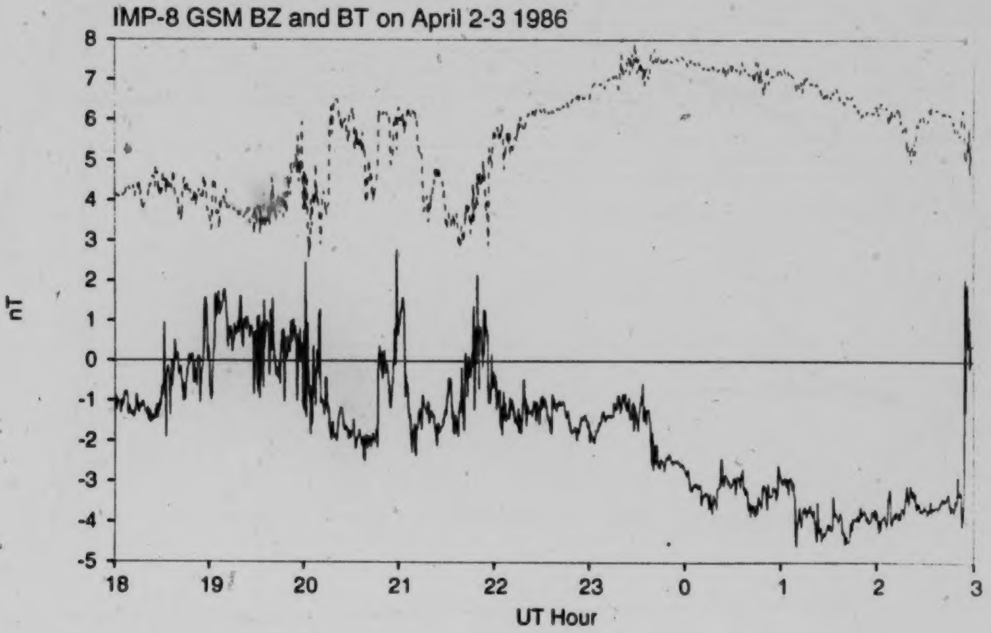
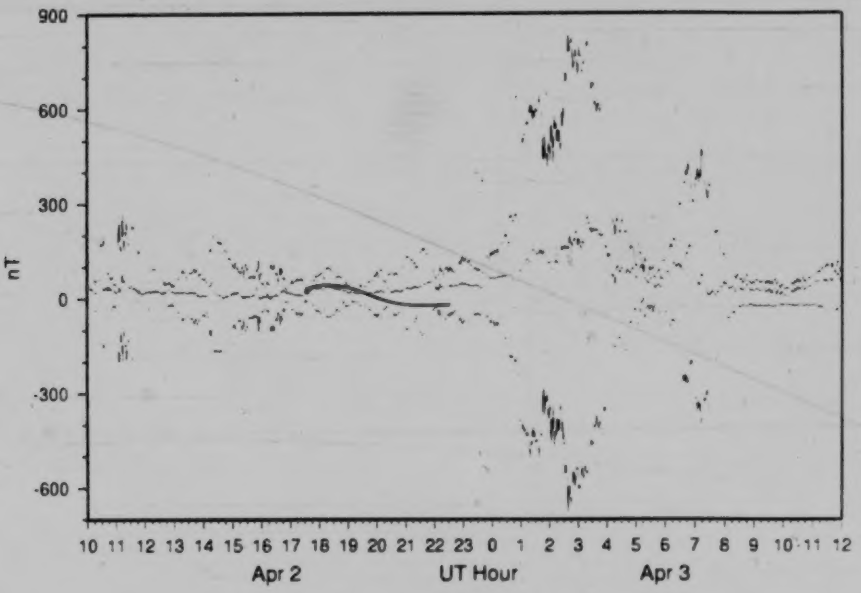


Figure 6.6 IMP 8 magnetic field on April 2-3 1986. GSM Z component (solid) and total field (dashed) are shown.



An overview of auroral zone magnetic activity is given by the AL and AU indices shown in Figure 6.7. The AL index decreases in response to the southward turning at about 2000 April 2. The strong southward turning about two hours later appears to not strongly enhance the AL index until it begins to decrease shortly before 0000 UT, possibly with a small expansive phase onset (very sharp drop) at about 0030. A major expansive phase is marked by the sharp decrease at 0050, after which activity remains strong, and is enhanced yet more at 0220. With the northward turning of the IMF, activity gradually decreases between about 0300 and 0500 UT April 3. The AU index increases slightly after AL for the two major onsets, and similarly decreases after 0300. Another episode of activity occurs between about 0600 and 0730. Since IMP-8 data ceased to be available at 0305 UT, the state of IMF is not known at this time.

Figure 6.7 Auroral Electrojet Indices on April 2-3 1986. Traces show AU (middle), AE (upper) and AL (lower). Time axis subinterval is 15 minutes.



It is worth noting that activity normally associated with substorm expansive phases (H bays at high latitude, Pi 2 activity, energetic particle injection in the equatorial plane) occurs in the late UT hours of April 2 without being reflected in any significant way by the AL or AU (and thus not by the AE) indices. Other aspects will be discussed later, but Figures 6.8 and 6.9 illustrate evening sector activity at Greenland stations at invariant latitude near 75°. After about 1600 UT April 2, various forms of activity are seen at these stations which are not reflected in the auroral electrojet indices. At the west coast station of Godthåb (GHB) positive X perturbations of up to about 100 nT are seen in the period 1600-2200 UT, followed by negative X perturbations. Comparable positive X is seen at the east coast station Kap Tobin (SCO) until about 1900, after which negative X prevails, including two abrupt onsets, part of the latter of which is also seen at GHB.

Figure 6.8 X Component Perturbations at Greenland stations on April 2 1986

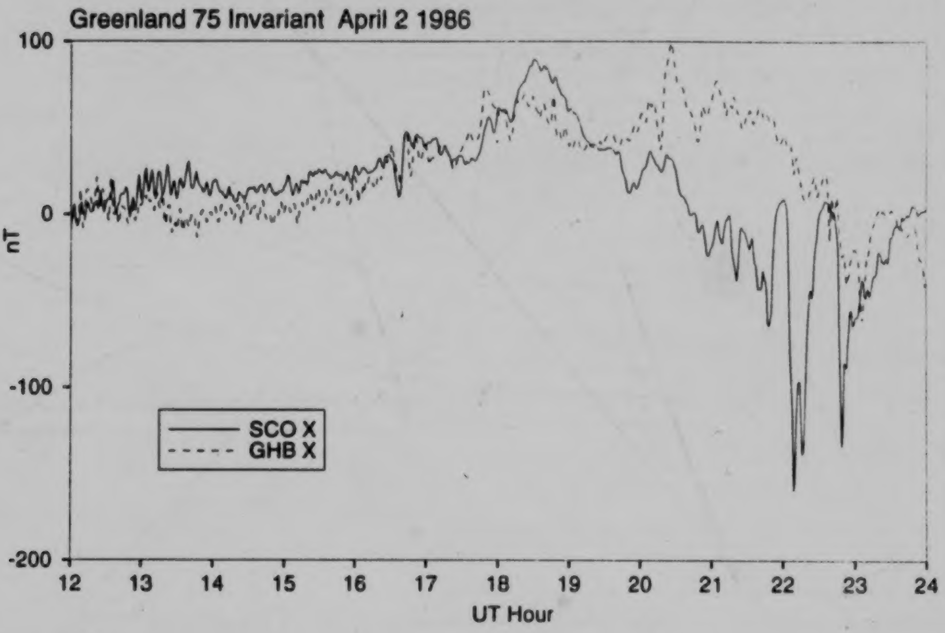
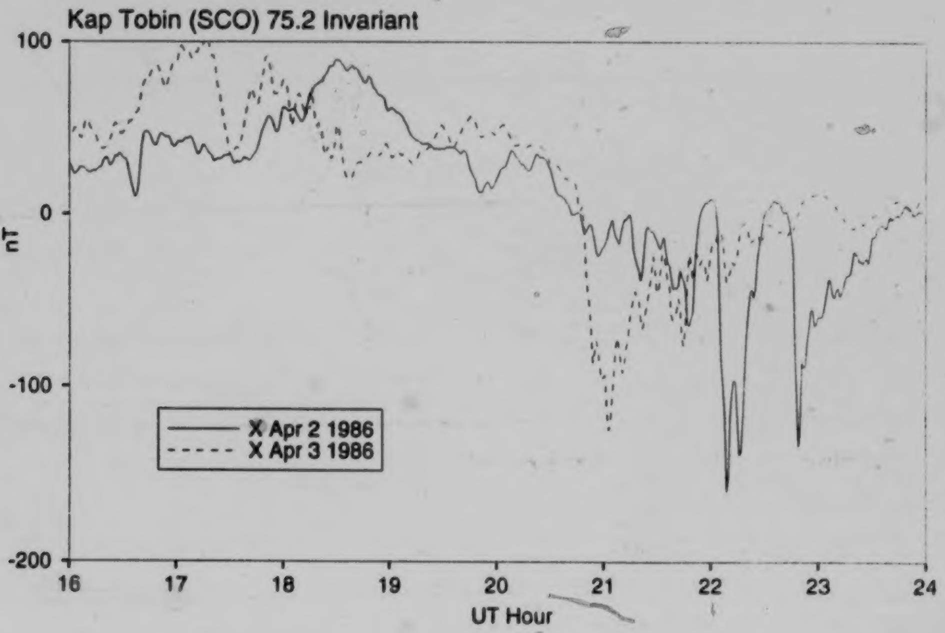
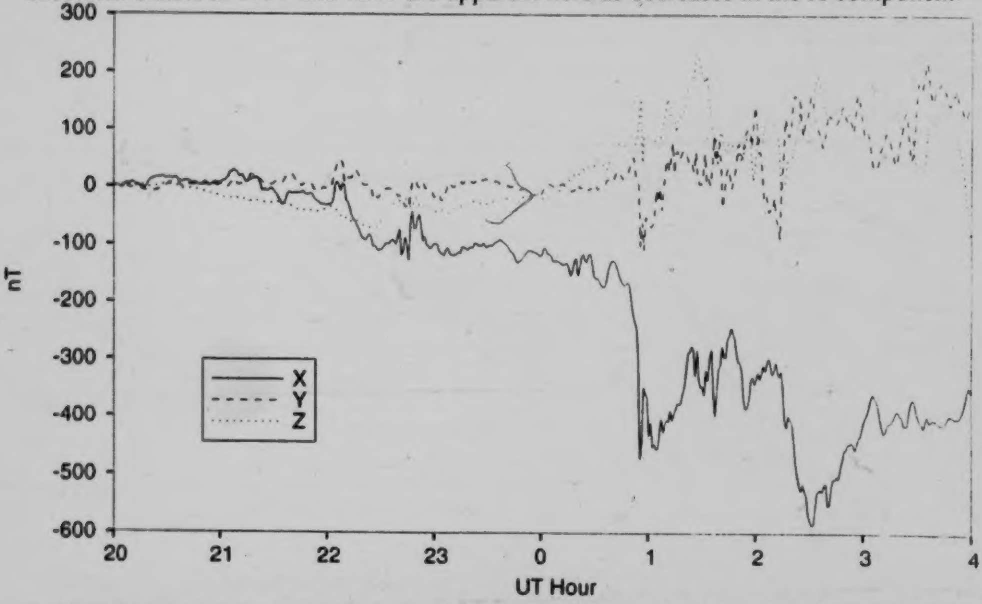


Figure 6.9 X Component Perturbations at Kap Tobin on April 2 1986, with quieter April 3 values shown for same UT.



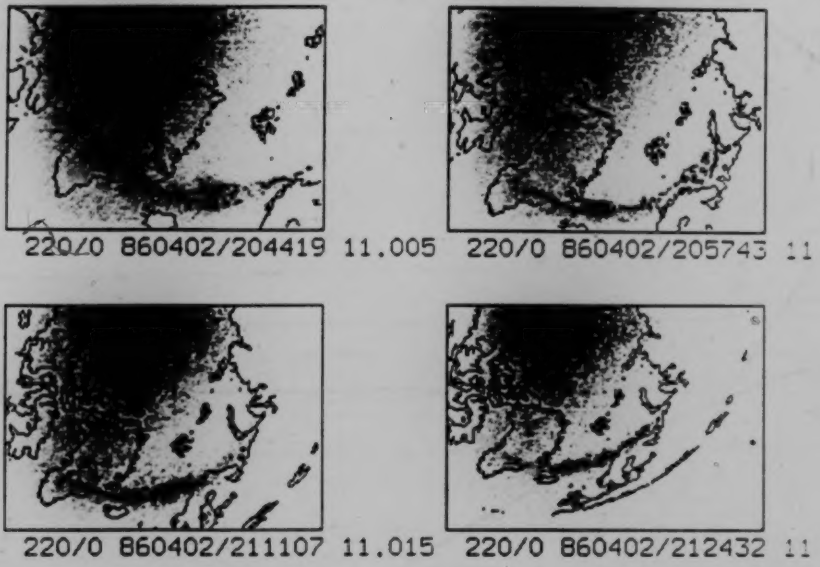
The negative X-component perturbations at GHB exceed 125 nT in magnitude, as may be seen by comparing a quieter (although perhaps not 'quiet') period on April 3 (Figure 6.9). The temporal and spatial pattern shown is consistent with the presence of an eastward electrojet (positive perturbations) while these stations are in the evening sector and later to a westward electrojet (negative perturbations). The latter appears to be linked to expansive phase activity in the midnight and morning sectors. The bias of the widely-used AE index to stations at about 65° geomagnetic latitude is well known [Baumjohann, 1986] and this is a dramatic example. Even onsets which are capable of producing well over one hundred nT of X perturbation at northerly stations have escaped incorporation into the standard auroral electrojet indices for this period. With this in mind it is useful to examine perturbations at Isafjordur, the most northerly station available in Iceland and the closest to Kap Tobin, shown in Figure 6.10. While both onsets have distinctive signatures at this station, primarily in the X component, they are small signatures of positive sign. This would indicate that the wedges involved are rather small in east-west extent, which is reinforced by the fact that only the second extends across Greenland, and in indistinct form as seen at the west coast (compare SCO and GHB signatures in Figure 6.8). Baumjohann [1986] also cites spacing of the 12 standard AE observatories in longitude as a reason for 'missing' activity, so it is conceivable that these small wedges would not be incorporated into the index even had they occurred at a more southerly latitude.

Figure 6.10 XYZ Perturbations at Isafjordur, Iceland, April 2-3 1986. X component negative bays of Kap Tobin at 2200 and 2245 here appear as positive bays. Major substorm onsets at 0055 and 0215 are apparent here as decreases in the X component.



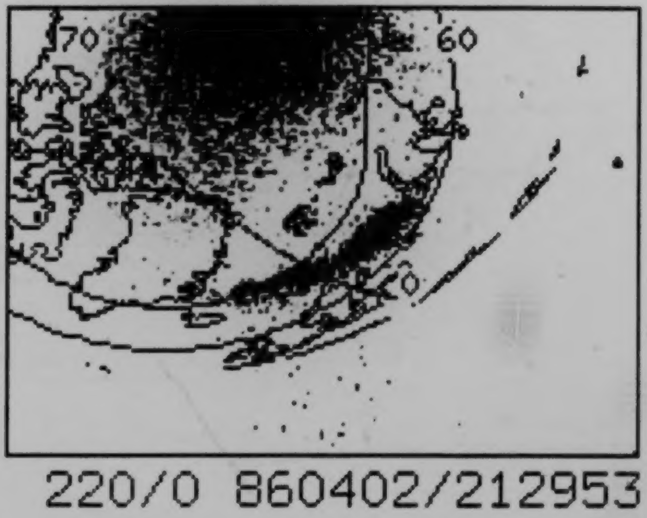
During this period the 75-station AE index calculated by Kamide [B. Emery, personal communication, 1995], incorporating these stations, is approximately 300, rather than the much lower level of the standard 12-station index (roughly 100). Due to the difficulty in obtaining sufficient magnetic data, indices computed from 50 or more stations are not commonly used [Baumjohann, 1986], but previous studies have shown, as here, that the ratio of AE(12) to AE(N), N>50, becomes small once AE is smaller than about 250 nT. The implication from results shown here and from such studies is clearly that activity at high latitudes is more common than AE(12) would lead one to believe. As is seen here, such activity can have signatures quite typical of expansive phase onsets.

Figure 6.11 Viking images in the vicinity of pre-onset Greenland by Camera 0 (LBH). Initial image (upper left) shows clearly a sun-aligned arc near the base of which activity occurs. Other images show isolated spots of increased emission (darker) within a discrete auroral zone.



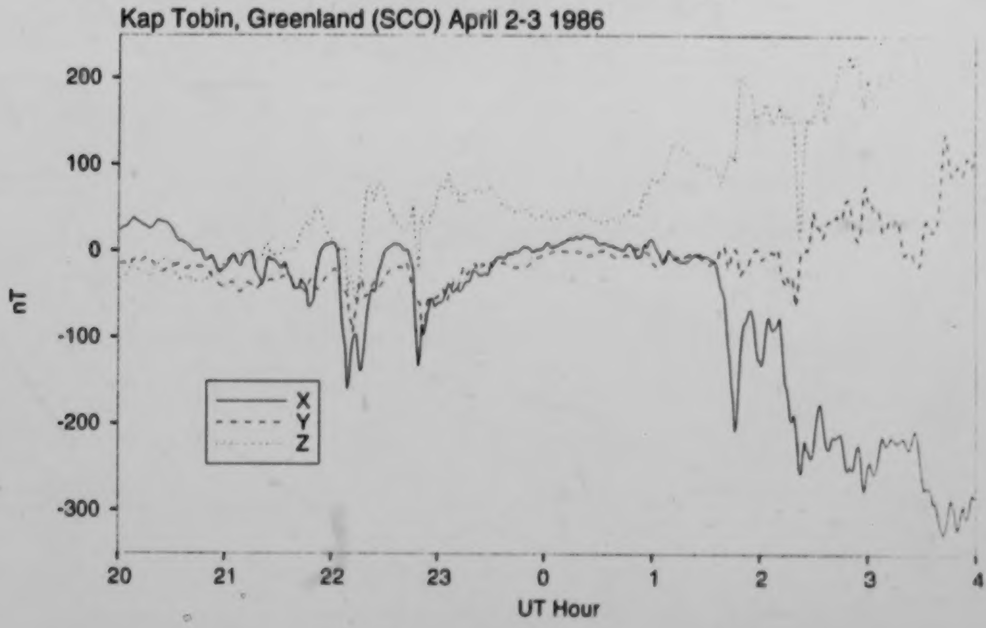
Further discussion of activity in the late UT day of April 2 is aided by examination of Viking UV images. Figure 6.11 provides these for the period immediately before the 2200 onset as seen at Kap Tobin. Images from this orbit (220) are not available at the time of onset but these images correspond to a period of disturbed westward electrojet as indicated by the negative X perturbations, varying in strength, at the east coast of Greenland. The auroral emissions are concentrated in a narrow band of variable intensity, except for a sun-aligned arc visible in the first image from 2044:19. In this image an intensified region with discrete features is seen over central Greenland, which is consistent with the interpretation of magnetograms from the same time (see Figure 6.8) as indicating converging electric currents, eastward over the west coast, westward over the east coast. These presumably diverge through upward field-aligned current, and if carried by precipitating electrons, enhanced UV brightness would be the result. This region appears less bright in subsequent images despite little change in the currents as indicated by magnetic signatures. Intensifications occur within the narrow band in subsequent images, yet there appears to be only one band of emission. The arrival of the intensified spots at the coast may correspond to small negative X bays in the magnetogram as seen in Figure 6.9 at 2115, associated likely with the spot which is near the east coast in the image taken at 2111:07. These features thus represent an active westward electrojet. Another significant aspect of these images is the orientation of the bright area corresponding to the westward electrojet. In Figure 6.11 this is seen to extend from northern Scandinavia to the central Greenland coast, which is about a 5° increase in latitude. Figure 6.12 shows EDFL latitude lines to illustrate this more clearly, and to show that the total northerly shift in latitude of the aurora from northern Russia to Greenland is in fact about 10°.

Figure 6.12 Contrast-enhanced Viking image showing tilted midnight sector electrojet. EDFL latitude lines at 60° and 70° are shown as is the EDFL geomagnetic midnight local time meridian.



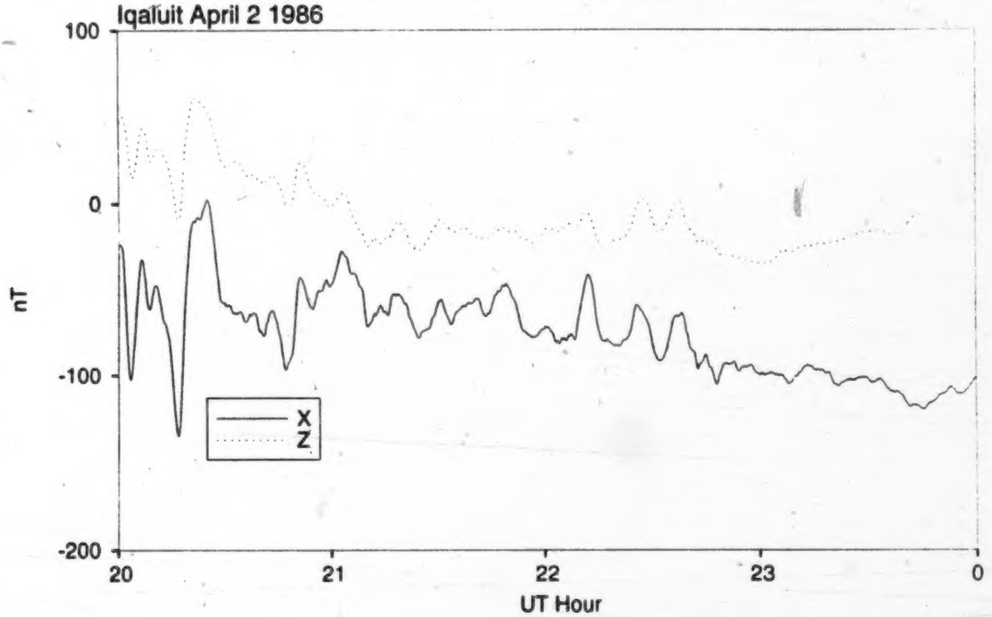
The latitudinal width of the bright region is not large, perhaps three degrees at widest. The narrowness of the electrojet is confirmed by the three-component Kap Tobin magnetogram of Figure 6.13. X component bays are accompanied by large changes in the Z component, which passes through zero at minima of X. While this behaviour is typical of any east-west aligned current system which has north-south motion, the rapidity of the variations in the Z component seems to indicate narrow electrojet width. The lack of significant X perturbations and only small Y and Z perturbations at station DNB (Danmarkshavn) 2.5° north of Kap Tobin (data not shown) bears out this conclusion. Viking images are not available at the times of high-latitude onsets. However, such high latitude activity on a northward-projecting spur of the electrojet (as inferred from emissions) has been seen in the declining stage of activity and inferred to be associated with the plasma sheet boundary layer far downtail [Henderson and Murphree, 1995]. The local time, roughly 22 hours, is similar to that in this case, but here the features are seen before a major onset rather than after. Yet another example of such a high-latitude evening sector spur of emission is illustrated by Elphinstone *et al.* [1993] with a Viking image from orbit 308 on April 18, 1986. An image in the same article shows that at 2000 UT April 2, that is to say at the start of the event under discussion here, the high-latitude spur was absent, having not yet formed.

Figure 6.13 XYZ magnetogram for Kap Tobin on April 2/3 1986. Large variations in the Z component accompany negative X bays before midnight.



Another interesting aspect of the intrusion of westward electrojet to high latitudes in the evening sector is an apparent relationship of the 2245 onset to cessation of pulsations. Pulsations with a period of roughly 10 minutes are seen in the Iqaluit magnetogram of Figure 6.14 to be pronounced from about 2200 to 2245 at this station. Their cessation and a decrease in the H component to a steady negative level appear to coincide with the 2245 onset.

Figure 6.14 $X_M Z$ magnetogram for Iqaluit on April 2 1986. Long-period pulsations are seen in the X_M component until 2245, after which it decreases and remains steady.



Farther east along the active westward electrojet, its larger scale effects may be seen by examining the magnetogram from Soroya, in northernmost Scandinavia. After about 2000 April 2, the westward electrojet intensifies steadily but impulsively as shown in Figure 6.15. At about 2100 UT one such intensification occurs with decrease in the X component, little change in the Z component, and a positive D signature which may indicate upward field-aligned current north of the station. At 2130 UT there is a similar signature but this time with a shift of Z to negative values, possibly indicating a northward shift of the electrojet. The actual onsets seen in Greenland at times 2200 UT and 2245 UT do not seem to be reflected here in a major way. At Saint John's, directly south of the Greenland activity, these onsets are reflected in a subtle way. The Y component at Saint John's, as seen in Figure 6.16, increases at 2200 UT April 2 in association with a brief negative X bay, which may be the signature of the Greenland onset seen from slightly to the west. The Y shift suggests that increased upward field-aligned current is flowing to the north of this station, and this shift remains approximately constant until a major onset occurs at 0055 UT April 3. This indicates that the onset activity in Greenland at 2200 UT April 2 brought the westward electrojet further west so that its upward field-aligned current could be detected at Saint John's.

Figure 6.15 XYZ magnetogram for Soroya on April 2 1986. From 2000 UT to 2200 UT the declining X component indicates growth in westward electrojet strength. Features at 2340 UT may be related to a pressure pulse in the solar wind.

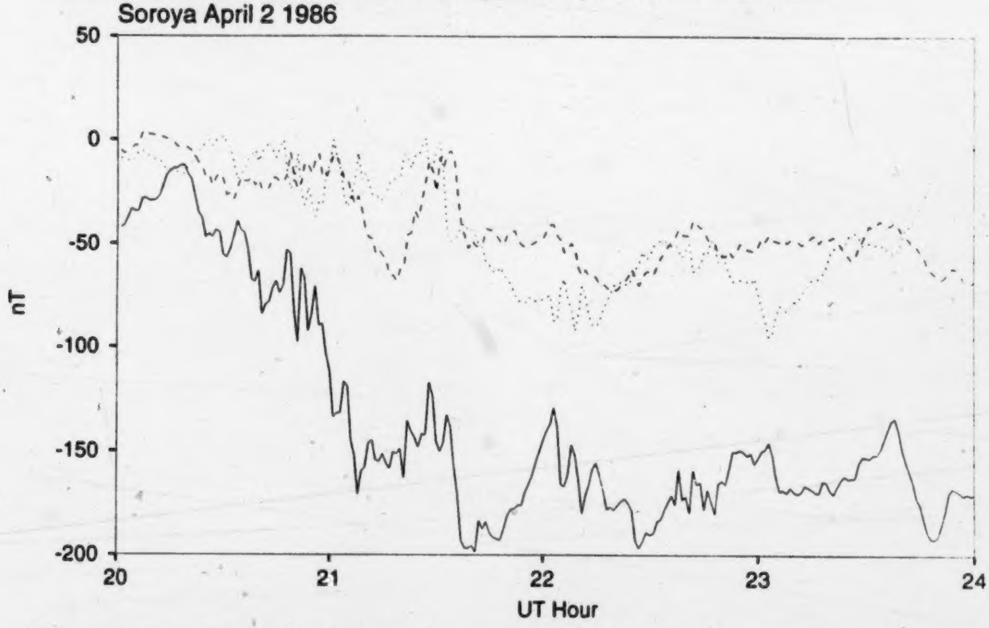
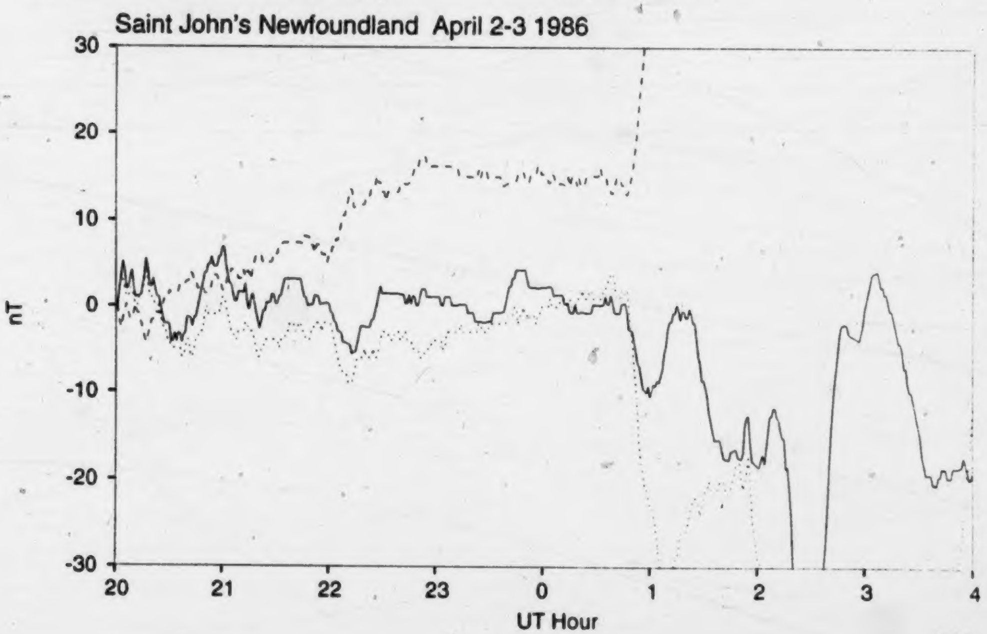
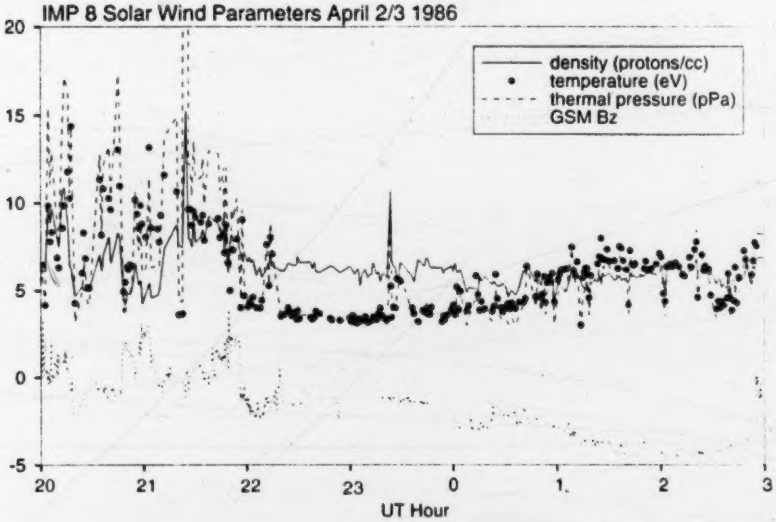


Figure 6.16 XYZ magnetogram for Saint John's for April 2-3 1986. Increase in Y at about 2200 UT may be associated with expansive phase activity in Greenland. Large perturbations accompany major expansive phase onsets at 0055 and 0215; scale chosen does not show these fully.

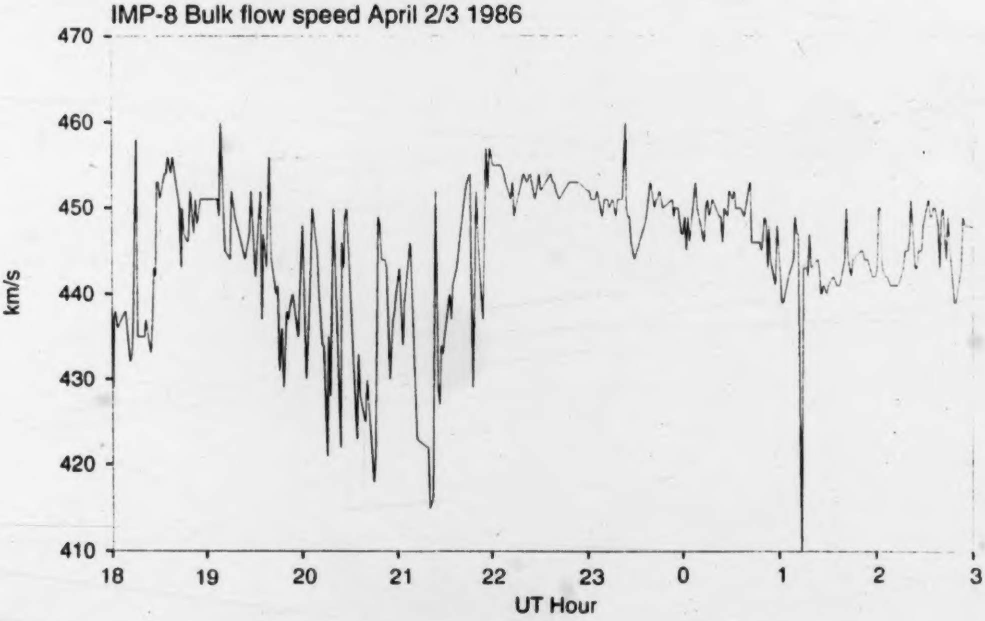


To give more context to the discussion of this preliminary period reference is again made to solar wind parameters as shown in Figure 6.17. From 2000 UT on April 2 until 0300 UT on April 3, the proton density was relatively constant. Temperature fluctuations before 2200 UT were large, but after the southward turning became smaller. Although the overall magnetospheric configuration responds to the bulk flow, reference back to Figure 1.1 shows that the bulk flow speed was quite constant at this time, and very close to the canonical value of 450 km/s. More details of the bulk flow speed are shown in Figure 6.18 at a more useful scale. It shows, prior to 2200 UT April 2, that the bulk flow speed was characterized by variations of about 30 km/s or 8% in a turbulent manner. The thermal pressure prior to 2200 UT also varied significantly, mainly due to temperature fluctuations. Both the ram pressure and the thermal pressure became much more constant after the southward turning. During this pre-southward-turning period of near-zero IMF B_Z, as noted, the westward electrojet grew stronger in a bursty fashion but it is hard to find actual onsets of activity. During this turbulent period it is tempting to suggest that conditions for sustained energy input to the magnetosphere did not exist for very long at any given time while abundant 'triggers' for release of any stored energy were present. A study of an interplanetary magnetic cloud and its effect on near-Earth space by Farrugia *et al.* [1993] showed high latitude onsets (not noted as such by the authors) prior to southward turning. At the time of southward turning, another such onset seemed to be triggered almost immediately, and the growth phase started as indicated by the eastward electrojet. Here, the electrojet growth seemed to take place during a period of near-zero IMF B_Z and it is only after the southward turning at 2200 UT that an onset occurred; indeed it came very shortly after southward turning. Since IMP-8 was radially downstream in the solar wind flow, it may be argued that the southward turning as seen near Earth was slightly earlier than that seen in the figure; it may even be that the onset at 2200 UT is associated with the brief northward turning and associated thermal pressure pulse seen at IMP-8 at approximately 2215. The other distinct onset at 2245 UT on April 2 may be associated with the very brief magnetic spike at about that time (see GSM B_Z in Figure 6.6 and also Y component in Figure 6.5). It is of course true that other quantities are varying in the solar wind, and pressure effects such as those illustrated in figure 6.17 should not be ignored in discussing trigger mechanisms. These aspects are not, however, further discussed here.

Figure 6.17 IMP-8 solar wind parameters on April 2-3 1986. Proton thermal pressure in pPa coincides numerically with the temperature for most values of the proton density. GSM B_z is once more shown for reference.



6.18 IMP-8 bulk flow speed on April 2-3 1986.



Finally, it is useful to consider to what extent the use of the term 'onset' above in relation to X bays at high latitude is justified. Certain phenomena are associated with an expansive phase onset that may not be characteristic of electrojet variations, which have also been inferred to occur based on the magnetic data (for example at Soroya). The availability of 2-second magnetic data from the Antarctic station Syowa allows spectral analysis to be performed. This station is conjugate to Iceland, and it has been shown that good conjugacy existed at this time (see section 3.d and Figure 3.9). Figure 6.19 shows the presence of Pi 2 pulsations at the times of both features mentioned as 'onsets'. Higher frequency Pi 1 or Pc 2 band activity was also seen, most prominently for the 2245 UT event (Figure 6.20). Activity in both of these pulsation bands is associated is a characteristic fashion with the substorm expansive phase (see Samson [1991], section 10.4.2). A second characteristic of true expansive phase activity is injection of energetic particles into the equatorial plane, but expansive phase activity may occur without this being detectable by spacecraft at any particular location [Kamide and McIlwain, 1974]. This may be understood by considering mapping. For these events in particular, where the activity seems to be narrowly localized at about 75° invariant latitude, the dipolar mapping would be to about 15 R_e at the equatorial plane, and the available synchronous satellites are considerably nearer to Earth than this and might be expected not to see the injections. This is even more the case since injections in near-dipolar tail geometry at about 15 R_e would result in dipolar drift paths which would quickly cross the magnetopause. Nevertheless, electron signatures which can be linked to these onsets were seen east of the likely injection region. Figure 6.21 shows electron data from the Los Alamos 1984-037A satellite, with small increases in the lowest energy channels at about 2250 UT and midnight. Most of these electrons were of about 30 keV energy, with drift speed eastward of about 20 km/s or 1.6°/minute at equatorial orbit. If particles causing the distinct 2250 UT signature are considered to have been injected at 2200 UT, then their longitudinal drift during this time would have been about 80°. Since the spacecraft was at geographic longitude 70°, this means that an injection at that time must have been at about 10° W geographic, slightly east of Greenland. Reeves *et al.* [1991] have determined that electron injections occur within an 'injection region' of limited longitudinal extent. The occurrence of a substorm current wedge in about the location needed to create the X-component bays seen at Kap-Tobin at 2200 April 2 is entirely consistent with the longitude inferred by roughly considering the drifts and with the known configuration of the injection region and its relationship to a current wedge. Examination of auroral electrojet indices (not shown) for an event studied by Reeves *et al.* [1991] (which occurred at 4.9 UT on October 16, 1983) revealed a similar location of the injection region with respect to the onset of the event. Similar considerations hold for the more dispersed signature at midnight in the event under discussion here. Thus these high latitude events do have the usual characteristics of an expansion phase onset: what lacks is characteristic eastward component signatures of the current wedge from lower latitude stations. This may be plausibly explained by the lack of stations in suitable positions and an inferred small scale size of the high latitude substorm current wedge. However, it has been shown above that perturbations were seen relatively near the onsets which are those of the interior of a current wedge (+X at Isafjordur). Thus small-scale wedges may have been present.

Figure 6.19 Syowa Pi 2 pulsations on April 2/3 1986. Onsets at 2200 and 2245 April 2 are detected, as are onsets on April 3 at 0055, 0145, and possibly 0215 UT

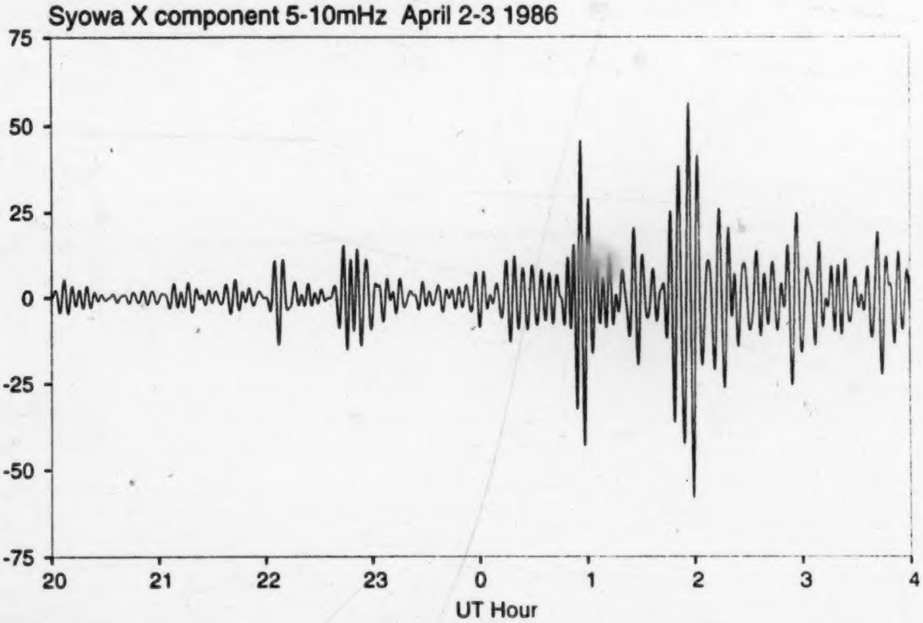

Syowa X component 5-10mHz April 2-3 1986

Figure 6.20 Syowa Pi 1/Pc 2 pulsations on April 2/3 1986. Onsets at 2200 and more prominently at 2245 April 2 are detected. The large onset at 0055 April 3 marks the start of quasi-continuous signal in this band.

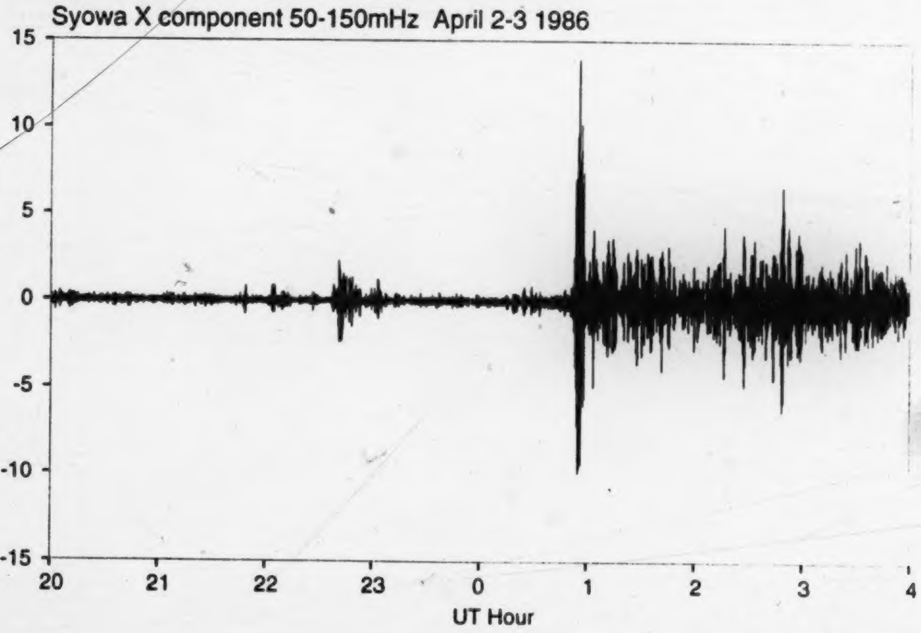
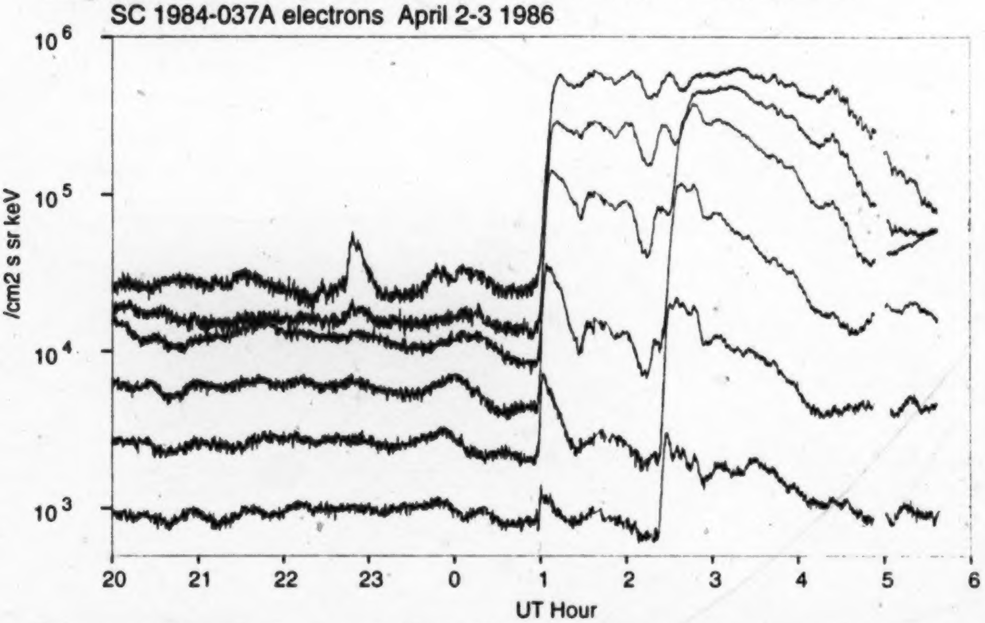
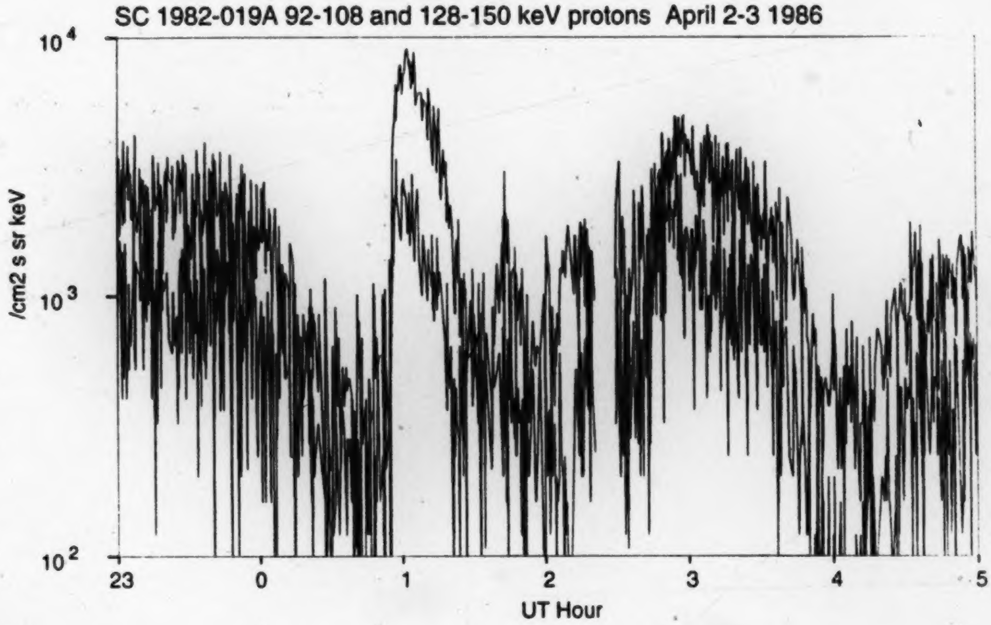

Syowa X component 50-150mHz April 2-3 1986

Figure 6.21 Electron flux at the SC 1984-037A synchronous satellite. Flux increases mark onsets at 2200 and 2245 April 2, and 0055 and 0215 on April 3. keV energies shown, from top: 30-45, 45-65, 65-95, 95-140, 140-200, 200-300.



The growth phase signatures associated with the major onset, which eventually took place at 0055 April 3 (as already noted in magnetic and particle data above), began at about 2245, as inspection of Figure 6.7 shows. There was an increase in the AU index and more notably a decrease of the AL index in characteristic fashion. It can be noted that during the period before 0215 the AU index never in fact became very large and this will be seen in modelling results to have characterized this event. The AU index is often associated with the evening sector 'driven system' eastward electrojet, and this is found below to be rather weak by comparison to the westward electrojet of the morning sector. Particle flux decreases in the equatorial plane near the midnight sector are also characteristic of the growth phase as canonically observed at that location [Savaud and Winckler, 1980]. The proton signatures, commencing at about 2345 as shown in Figure 6.22, are characterized by flux decreases of at least a factor of 10 until the injection accompanying expansion phase onset. An increasing tail current results in the satellite sampling mirroring particles from greater radial distances; these are less numerous and thus the particle flux at the spacecraft decreases during the growth phase. Equivalently, in the terms of Savaud and Winckler, the outer radiation zone is displaced inward toward the Earth, and the flux decrease observed is the zone's normal negative flux gradient with radial distance. Data from this event were presented in Chapter 1 (Figure 1.22) to illustrate dipolarization at geosynchronous orbit. The increasing deviation of the H (northward) component from quiet time values indicated growth of the tail current. That current caused the nightside field to become more 'tail-like' with a smaller magnitude of the northward component and a larger magnitude of the radial component, as seen in the Z (radially inward) component in the figure.

Figure 6.22 Proton flux at the SC 1982-019A synchronous satellite. Flux dropout gradually occurs from 2345 UT April 2 to the time of injection at 0055 UT April 3. That injection is not dispersed. A much smaller injection occurs at about 0130. Peak at about 0300 is interpreted as a dispersed signature from 0215 injection which occurred west of the spacecraft so that protons had to drift almost completely around the Earth before detection. Fluxes are shown for 92-108 keV (top) and 128-150 keV energies.

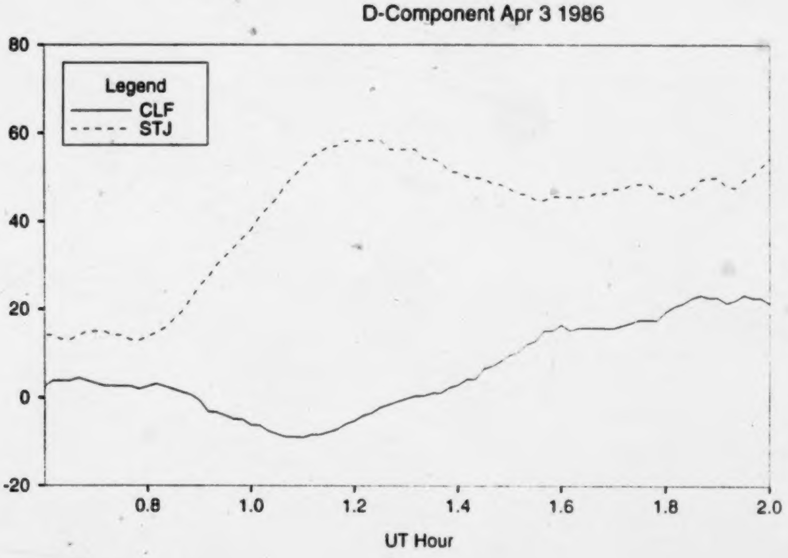


Global modelling becomes possible at 0030 UT using data supplied by Barbara Emery [personal communication, 1995] for the period 00 to 03 UT. The restriction to times greater than 0030 UT is due to data dropouts from the two CANOPUS stations operational, but undergoing testing, at this date. These stations are essential for delineating the evening sector eastward electrojet which, as mentioned above, is rather weak. Skone [1994] used latitude profiles from these and nearby stations to conclude that the evening sector electrojet moved from being centred at 70° (in fact Churchill at 68.5°) to 66° (Gillam) between 0040 and 0050 on April 3. Thus accurate modelling of the evening sector may contain useful information about the growth phase of this event.

d. First major onset

After approximately one hour of growth phase, a major substorm onset was observed by multiple methods, starting at about 0050 UT April 3. Classic signatures of the formation of a current wedge are the midlatitude Y_M component perturbations on each side of the Atlantic shown in Figure 6.23, steep X_M component decreases at appropriately situated auroral zone stations as seen in Figure 6.24, Pi 2 pulsations in the 5-10 mHz band as seen in Figure 6.19, and higher frequency Pi 1/Pc 2 pulsations as seen in Figure 6.20. The Y_M component perturbations change slope at 0050, and there is some change in Pi 2 activity at this time. Pi 1 band pulsations increased markedly in amplitude at 0055, concurrent with auroral zone bays.

Figure 6.23 Perturbations at Chambon-La-Forêt, France, and St. John's, Newfoundland, Canada, showing appearance of Y_M-component bays of opposite polarity.



Auroral station distribution is not optimal, however, as the current wedge straddles the Atlantic. While the Y_M perturbations clearly indicate field aligned current downward over the eastern Atlantic shore and upward over the western shore, it is not possible to observe mid-latitude perturbations under most of the current wedge. At auroral latitudes, onset timing is also not very precise due to the separation of stations. From Iceland and Greenland, the onset appeared earliest, at 0055 (see Figure 6.24). The dispersionless proton injection seen in Figure 6.22 also occurred at 0055. Thus 0055 is used as the nominal onset time although there are indications of lesser activity a few minutes earlier. All available Northern Hemisphere stations were used in modelling this event. Perturbations with a temporal spacing of five minutes, with quiet day April 4 removed, were supplied by Barbara Emery. Comparison with CDAW data confirmed that these data corresponded extremely well to CDAW data at available stations, provided that 2000 April 2 was subtracted off as a 'quiet' baseline. The exception was Dixon, where at that reference time perturbations in the H component of up to about 75 nT were occurring. Stations used are shown in Figure 6.25 and the stations are listed in Table 6.8. Figure 6.26 has the same orientation as does Figure 6.25 and shows initial conditions, details of which are discussed below. It is clear that there is, as stated above, a lack of midlatitude and subauroral stations due to the presence of the Atlantic Ocean. This may be a limiting factor in the accuracy of modelling this phase of activity. The D component perturbations clearly show that a substorm current wedge was present across the Atlantic until about 0135 (decimal hour 1.6 in Figure 6.23). Complementing the ground data, a DMSP F7 spacecraft pass through or near the substorm current wedge during this time interval allows modelling of Pedersen currents which may not have a strong ground signature. A similar study, with simpler magnetic modelling and limited use of ground stations, but with detailed analysis of DMSP particle data, was performed for a substorm current wedge of April 25, 1985 by Lopez *et al.* [1991]. Here the focus is reversed, being primarily on ground data.

Figure 6.24 X_m component perturbations at Narssarssuaq, Greenland, and Husafell, Iceland. These are one-minute data with zero reference point at 2000 April 2. Steep X_m component decrease seen at about 0055 UT at Husafell has an analogue at about 0102 UT at the further west station of Narssarssuaq.

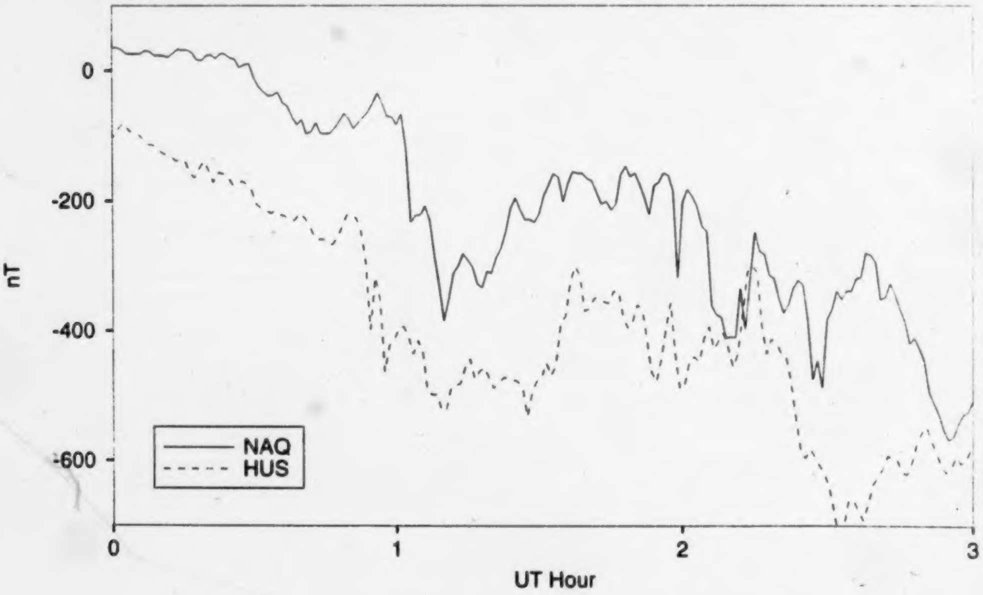


Table 6.8 Stations used in modelling for CDAW Event 9B. See Figures appropriate to each time period for stations used at any particular time.

Code	Name	Geodetic Latitude	Geod. Long.	CD Latitude	CD Longitude	CD Dec.
THL	Thule	77.5	290.8	88.6	15.5	-13.8
ALE	Alert	82.5	297.5	86.1	163.6	-155
RBC	Resolute Bay	74.7	265.1	83.2	-65.0	41.6
NRD	Nordkapp	81.6	343.3	80.8	132.2	-78.6
UMQ	Umanaq	70.7	307.9	80.6	40.5	-22.3
MBC	Mould Bay	76.2	240.6	79.5	-100.3	52.9
GDH	Godhavn	69.3	306.5	79.4	35.0	-18.3
DMH	Danmarkshavn	76.7	341.4	79.2	105.7	-54.4
ATU	ATU	67.9	306.4	78.1	33.0	-16.3
DNB	Daneborg	74.2	339.9	77.8	94.8	-45.1
STF	Sondre Strømfjord	67.0	309.3	76.9	36.6	-17.2
CBB	Cambridge Bay	69.1	255.0	76.8	-61.3	28.4
SCO	Kap Tobin	70.5	338.0	75.3	82.5	-35.0
IQL	Iqaluit	63.8	-69.2	74.9	2.9	0
GHB	Godthåb	64.2	308.3	74.3	31.9	-13.6
BLC	Baker Lake	64.3	264.0	73.8	-41.1	17.1
FHB	Frederikshåb	62.0	310.3	72.0	33.3	-13.1
HIS	Heiss Island	80.6	58.0	71.5	156.5	-28.2
BJN	Bjørnøya	74.5	19.2	71.0	125.0	-36.4
ESC	Eskimo Point	61.1	266.0	70.9	-35.4	13.4
NAQ	Narssarssuaq	61.2	314.6	70.7	38.8	-14.6
LRV	Leirvogur	64.2	338.3	69.7	72.3	-25.0
YKC	Yellowknife	62.4	245.5	69.2	-63.7	22.0
BRW	Barrow	71.3	203.3	69.0	-116.6	32.6
FCC	Fort Churchill	58.6	265.9	68.6	-34.0	12.0
ZHE	Mys Zhelaniya	77.0	68.6	67.4	157.7	-19.0
SOR	Söröya	70.5	22.2	67.2	121.0	-29.8
CCS	Cape Chelyushkin	77.7	104.3	66.6	177.4	-2.3
ALT	Alta	69.9	23.0	66.4	120.7	-28.6
PDB	Poste de la Baleine	55.3	282.2	66.3	-9.8	3.3
GIL	Gillam	56.4	265.4	66.2	-33.5	11.1
KIL	Kilpisjarvii	69.0	20.7	66.1	118.1	-28.5
KEV	Kevo	69.8	27.0	65.7	123.7	-27.7
KAU	Kautokeino	69.0	23.0	65.7	119.8	-27.9
CMO	College	64.9	212.2	65.1	-101.0	26.5
MUO	Muonio	68.0	23.5	64.7	119.1	-26.8
MKM	Maliye Karmakuly	72.3	52.5	64.5	144.0	-22.0
BEY	Ostrov Belyj	73.3	70.0	63.7	155.8	-16.0

SOD	Sodankyla	67.4	26.6	63.6	120.9	-25.5
PEL	Pello	66.9	24.1	63.6	118.4	-25.7
MMK	Murmansk	68.3	33.1	63.4	126.6	-24.7
CPS	Cape Schmidt	68.9	180.5	63.3	-130.6	24.1
DIK	Dikson	73.5	80.6	63.2	162.5	-11.8
LER	Lerwick	60.1	-1.2	62.2	89.8	22.8
MEA	Meanook	54.6	246.7	62.0	-56.2	16.1
KHR	IZMIRAN KHR	71.1	66.8	61.9	152.5	-16.0
TAM	Tambei	71.5	71.8	61.8	156.0	-14.4
YSH	IZMIRAN YSH	69.8	60.8	61.2	147.6	-17.4
TIK	Tiksi	71.6	128.9	60.9	-167.3	7.7
SYH	IZMIRAN SYH	70.1	72.5	60.4	155.7	-13.5
SIT	Sitka	57.1	224.7	60.3	-82.1	20.6
GLL	Glenlea	49.6	262.9	59.3	-34.0	9.6
ARK	Arkangel'sk	64.6	40.5	58.9	129.4	-20.4
CKA	IZMIRAN CKA	68.5	73.6	58.8	155.8	-12.5
STJ	Saint John's	47.6	307.3	58.0	23.5	-6.6
ESK	Eskdalemuir	55.3	-3.2	58.0	84.1	19.5
LOV	Lovo	59.3	17.8	57.8	107.0	-21.3
YSL	Yar-Sale	66.8	70.8	57.4	153.1	-12.8
OTT	Ottawa	45.4	284.5	56.5	-5.8	1.6
VAL	Valencia	51.9	349.8	56.2	75.0	-17.6
LNN	Leningrad	60.0	30.7	56.1	118.5	-19.8
NEW	Newport	48.3	242.9	55.2	-57.4	14.2
VIC	Victoria	48.5	236.6	54.4	-64.5	15.3
HAD	Hartland	51.0	-4.5	54.2	80.5	17.6
MGD	Magadan	60.1	151.0	51.2	-147.9	11.9
POD	Podkammenniye	61.6	90.0	50.9	165.7	-5.6
FRD	Fredericksburg	38.2	282.6	49.3	-7.8	1.9
SVD	Sverdlovsk	56.4	58.6	48.5	139.9	-13.0
PET	Petsamo	52.9	158.4	45.0	-139.7	12.0
FRN	Fresno	37.1	240.3	43.9	-56.4	11.6
ODE	Odessa	46.8	30.9	43.5	112.5	-15.1
TOL	Toledo	39.9	356.0	43.4	76.3	-14.2
IRT	Irkutsk	52.2	104.4	41.1	176.2	-1.2
TUC	Tucson	32.3	249.2	40.5	-45.5	9.4
KGD	Karaganda	49.8	73.1	40.4	150.1	-8.6
DLR	Del Rio	29.9	259.1	39.4	-34.1	7.2
MMB	Memembetsu	43.9	144.2	34.5	-149.8	7.8
SJG	San Juan	18.1	293.9	29.2	5.2	-1.1
KAK	Kakioka	36.2	140.2	26.5	-152.3	6.4
HON	Honolulu	21.3	202.0	21.5	-91.5	12.0
KNY	Kanoya	31.4	130.9	21.0	-160.2	4.4

BAG	Baguo	16.4	120.6	5.5	-168.9	2.2
GUA	Guam	13.6	144.9	4.5	-145.2	6.5

Having established that a distinctive onset occurred at 0055 April 2, a digression into the question of solar wind triggering of onsets may be appropriate. Inspection of the IMF in Figure 6.6 will show that southward field had dominated for nearly three hours at the time of this major onset. That an onset would occur, in the standard view of accumulation of energy in the magnetotail which is then released at onset time, is not unexpected. For this onset there is no particular signal in the solar wind which can account in a clear way for its triggering. This will be seen to also be the case for the onset which occurs at 0215 April 3, discussed later in this chapter. In Chapter 7 the CDAW 9A event is discussed, which appears to conform to the more standard 'canonical substorm' featuring a southward turning, growth phase, northward turning and onset. That the events under discussion in this chapter do not conform to this model is not a great cause for concern as the question of external versus internal triggering of substorms has been actively discussed over at least the last decade. Farrugia *et al.* [1993] found little evidence of solar wind triggering during an extended southward IMF period which featured numerous substorms. A recent contribution to this discussion is the analysis of comprehensive data sets by Henderson and coworkers at Los Alamos in cooperation with Murphree at University of Calgary [1995]. Reanalysis of earlier data sets of McPherron showed that although no clear triggers of onset could be identified in the solar wind data, it was not possible to rule out triggering by small IMF fluctuations which are ubiquitous. The same might be said for the event considered here where there is an increase in the IMF B_Y component at about onset time. It is difficult to decide upon a threshold below which fluctuations no longer should be considered to be significant in terms of ability to trigger. This being the case, Henderson and coworkers attempted to identify onsets initially through Los Alamos synchronous particle detectors observing an injection having characteristics similar to those shown in Figure 6.21, that is, a large increase in particle flux which is near-dispersionless in the midnight sector and has dispersion away from midnight. In a qualitative sense it was required that the solar wind B_Y, B_Z and dynamic pressure be steady for 30 minutes before candidate non-triggered onsets. In a careful study, 10 onsets were found which satisfied these criteria and also had solar wind data available from a suitably-placed (noon sector) IMP-8. In all but one case, the solar wind was quite free of disturbances which could conceivably be claimed as triggers for onset. The results cited here, while not having solar wind data quite as free of potential triggers, could be added to this list, strengthening the claim of Henderson *et al.* that at least some substorms are internally triggered (more accurately one would have to say not convincingly externally triggered). It is also worth noting in Figure 6.7 that there is no increase in the magnitude of the auroral electrojet indices at the time of northward turning at 0300. While this and absence of further major injection activity (Figure 6.21) are not strong proof that another onset did not occur, there is no evidence to suggest that northward turning at 0300 did trigger further expansive phase activity.

Figure 6.25 Locations of stations used in CDAW 9B modelling. Nadir 60.70. Insets show codes of stations in Fennoscandia and in Greenland/Iceland.

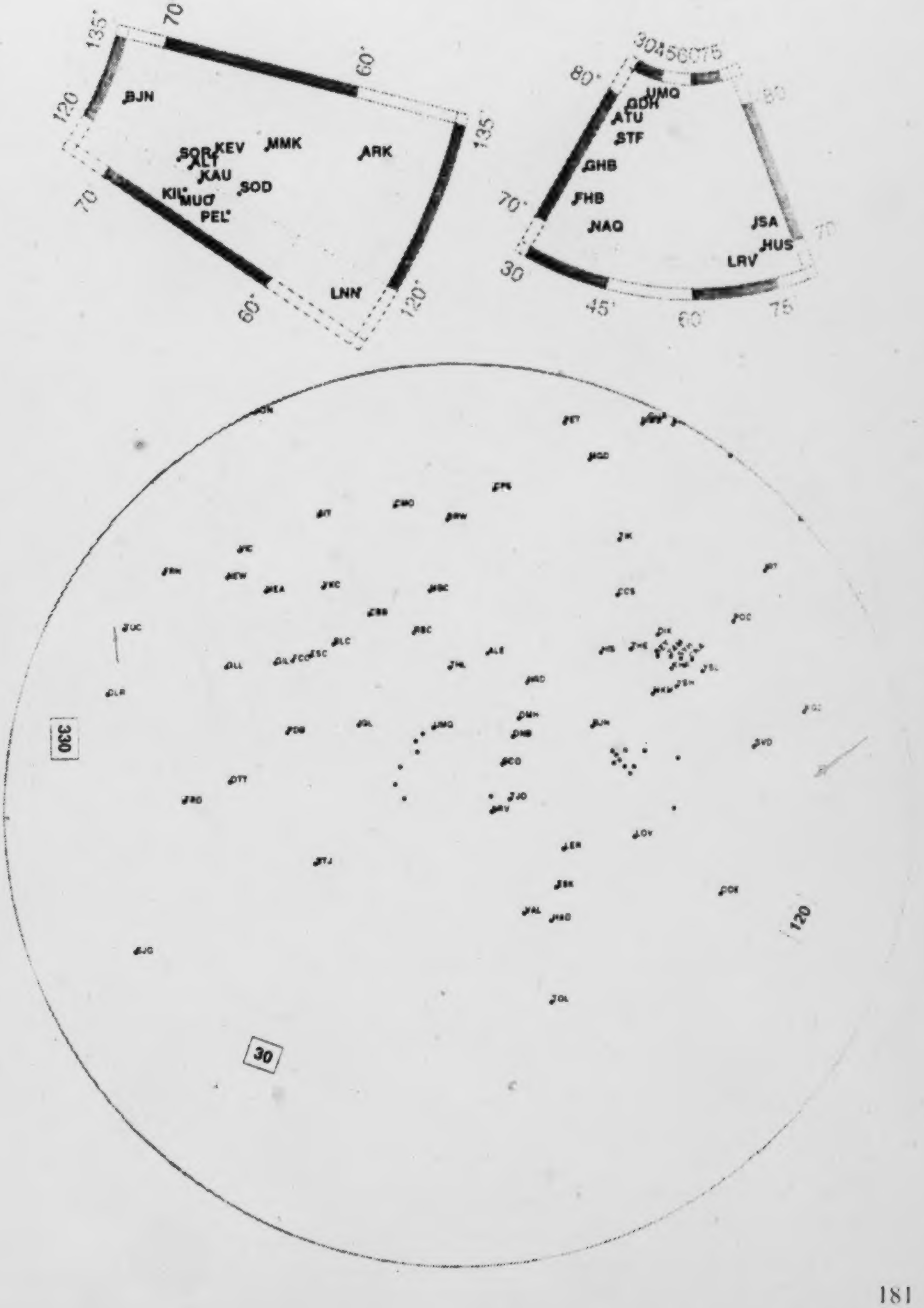
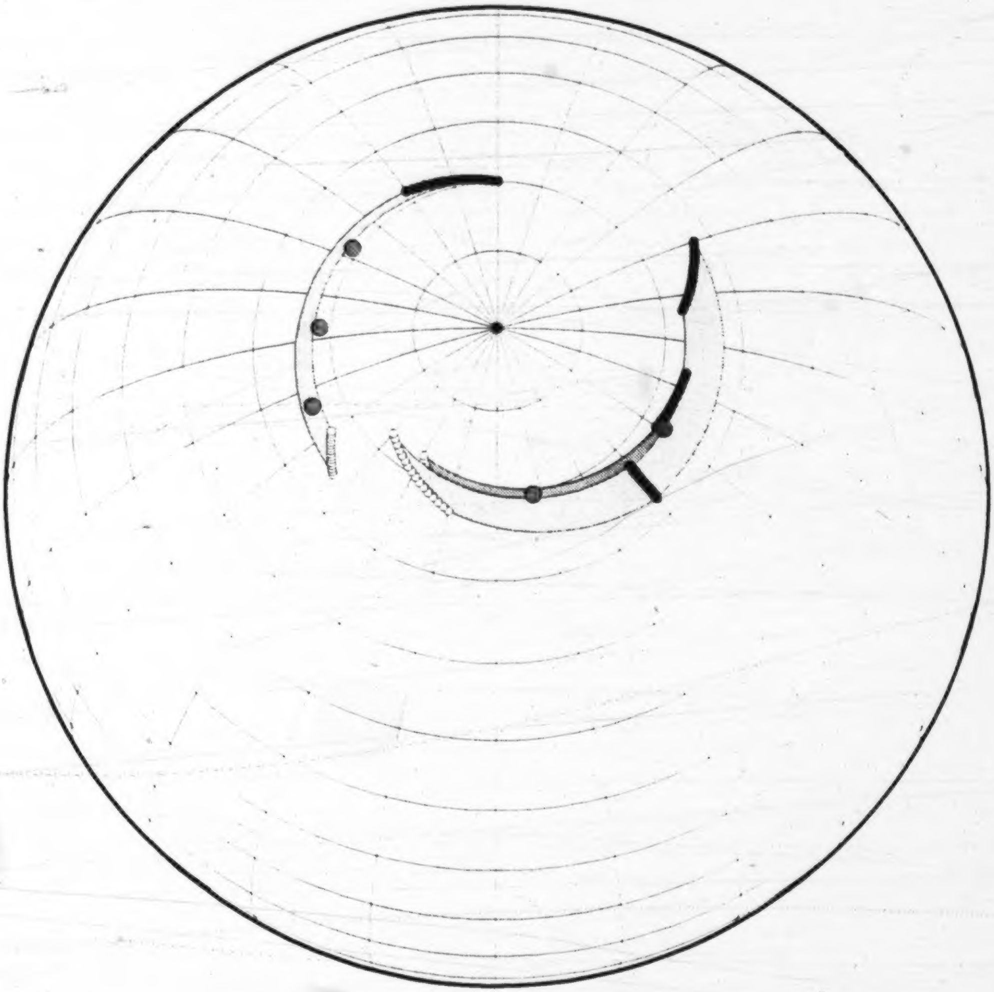
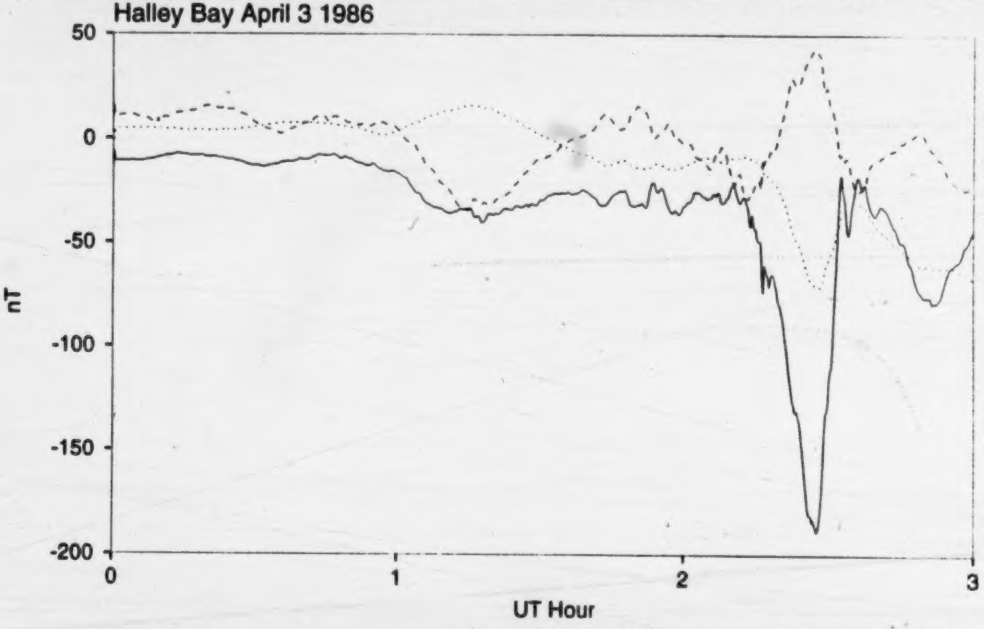

135°
70
60°
120
135°
BJN
SOR
KEV
ALT
MMK
ARK
KAU
SOD
KIL
MUO
PEL
70
60°
LNN
120°
30
45
60
75
80°
UMQ
GDH
ATU
STF
GHB
70°
FHB
NAQ
JSA
HUS
LRV
80
70
30
45°
60°
75°
330
30
120

Figure 6.26 Initial conditions used in modelling for 0100-0200 UT April 3 1986. Shading within current systems serves only to distinguish them. Downward current is indicated by filled black circles; upward current by white circles. Stations are shown for reference as grey circles, from east (clockwise): SOR, LRV, PDB, FCC, YKC.



Initial conditions shown in Figure 6.26 represent suitable initial conditions deduced from inspection of the ground perturbations and partly through previous modelling runs. The most notable aspect of the modelling system as shown is that it is necessary to break the westward electrojet into two segments joined at 105° magnetic longitude. The current is forced to be the same in each segment: there is no upward or downward net current flow at the junction point. The latitudes of the northern and southern borders at 105° are forced to be identical in each part of the joined system but are otherwise free to vary. The introduction of this joint was found to be needed in order to allow a good fit at stations under the westward electrojet. In other words, the westward electrojet in this case cannot be well represented by a single system with constant derivative of latitude versus longitude. Most particularly, the modelling indicated that an approximation in which the electrojet was at constant centred dipole latitude was not a good one. Such a poor forward model not only does not allow a good fit to be obtained in the region under the westward electrojet, but further contributes to making the problem ill-conditioned by dominating the weighted χ^2. The jointed nature of the electrojet as used in this forward model may appear unusual. It might, however, be expected on statistical grounds. Locations of the electrojet borders as deduced from the study of the Alberta chain rotating were shown in Figures 1.9 and 1.10. In the low activity case (Figure 1.9) the westward electrojet is found poleward of the 'evening sector' eastward electrojet, rising from about 70° at about 0200 MLT to more than 75° at 2000 MLT. This configuration is very similar to that found here through extensive preliminary modelling. At higher activity levels (Figure 1.10) the northward-projecting electrojet spur no longer seems to be present. Akasofu [1964] also noted high-latitude evening sector auroral luminosity (see Figure 1.19) which could quite plausibly be associated with the northward spur of the electrojet. Thus, a configuration similar to that required to obtain good fits seems to have been found independently in two other studies. It would be useful to have direct observational support for the contention that the jointed westward electrojet does not simply arise from a lack of data. For this reason, and since the presence of the Atlantic Ocean does not permit an optimal station placement, conjugate data from Halley Bay may be considered to determine whether perturbations arise which would be due to an electrojet at constant latitude. The raw Halley Bay perturbations are shown in Figure 6.27 and it must be recalled that transformations of magnetic components are required in order to compare data between hemispheres. The H component remains the same, negative for a westward electrojet. However, the Z component from the electrojet has reversed behaviour, being positive in the southern hemisphere data if the electrojet is poleward rather than equatorward of a station. It is seen that prior to 0100, there were no significant or systematic perturbations seen at Halley Bay. Around 0100 a negative X_{LM} bay indicates the presence of the westward electrojet, but the approximately equal positive Z perturbation shows that it is well poleward of the station. The Y_{LM} component indicates upward current flow poleward of the station, consistent with results from Saint John's placing this local time sector near the west end of the current wedge. There is no sign of eastward electrojet at any time. The gap between the eastward and westward electrojets in the initial model seems justified by this observation.

Figure 6.27 Magnetic perturbations at Halley Bay on April 2-3 1986. This station is conjugate to northern Newfoundland. Local magnetic (LM) system used is very close to centred dipole system at this location. X_{LM} shown solid, Y_{LM} dashed, Z dotted.



Iterations at each time step were started from the conditions shown in Figure 6.26 as discussed above. Final solutions differ since the perturbations at any given time vary and the routine optimizes the model's fit to them. The initial conditions shown do not correspond to a solution at any time modelled but are such as to allow solutions to be obtained during the time period studied. The current wedge in the initial model partially overlaps the westward electrojet: this is not claimed to be a feature expected in the solutions and indeed the solutions generally remove this overlap. The overlap is, rather, simply an arbitrary choice. The ability to start from a condition like this and arrive at solutions which do not feature overlap should be seen as a strength of the method. The placement of the evening sector eastward electrojet does correspond fairly much to what will be seen in the solutions. The fact that it stops at roughly 0° geomagnetic longitude is difficult to ascertain from the stations included in the modelling. As already alluded to, use of Halley Bay conjugate data provides support for the lack of further extension of the eastward electrojet into the morning sector. There is nothing in the modelling data set to strongly constrain the easternmost extent of the eastward electrojet, and the solutions do not deviate strongly from the initial condition due to this lack of constraint. From the modelling there would be no strong reason to make inferences about where the eastward electrojet ends: the conjugate data, however, suggest that the configuration used in the initial conditions, and largely unchanged by the modelling, is a reflection of reality. The separation of eastward and westward electrojets is a noteworthy result consistent with all available magnetic data. Modelling results at five minute intervals are shown in Figures 6.28 through 6.34 for the period from 0055 to 0125, and these results are discussed on the next text page.

Figure 6.28 Modelling results at 0055 UT April 3 1986.

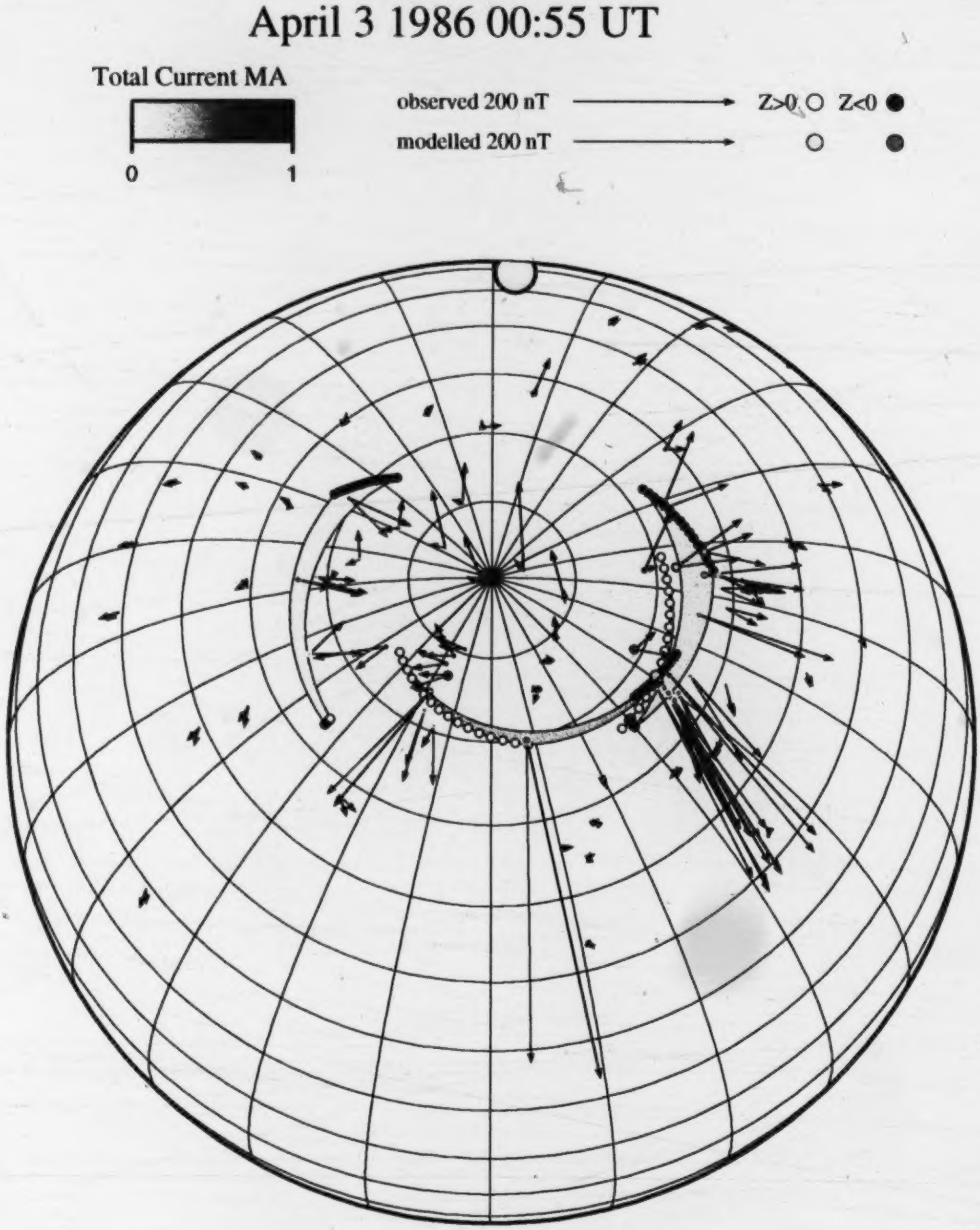

April 3 1986 00:55 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.29 Modelling results at 0100 UT April 3 1986.

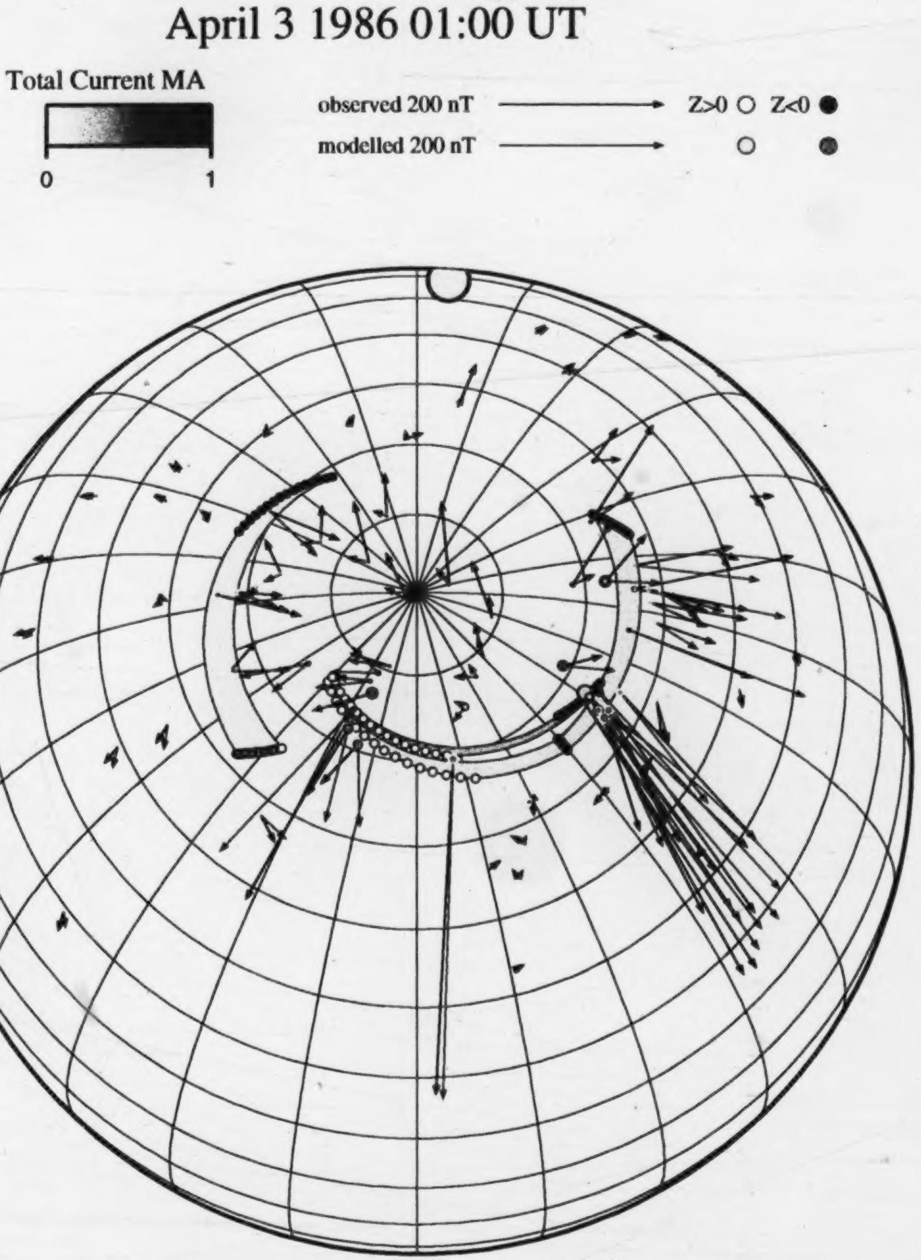


Figure 6.30 Modelling results at 0105 UT April 3 1986.

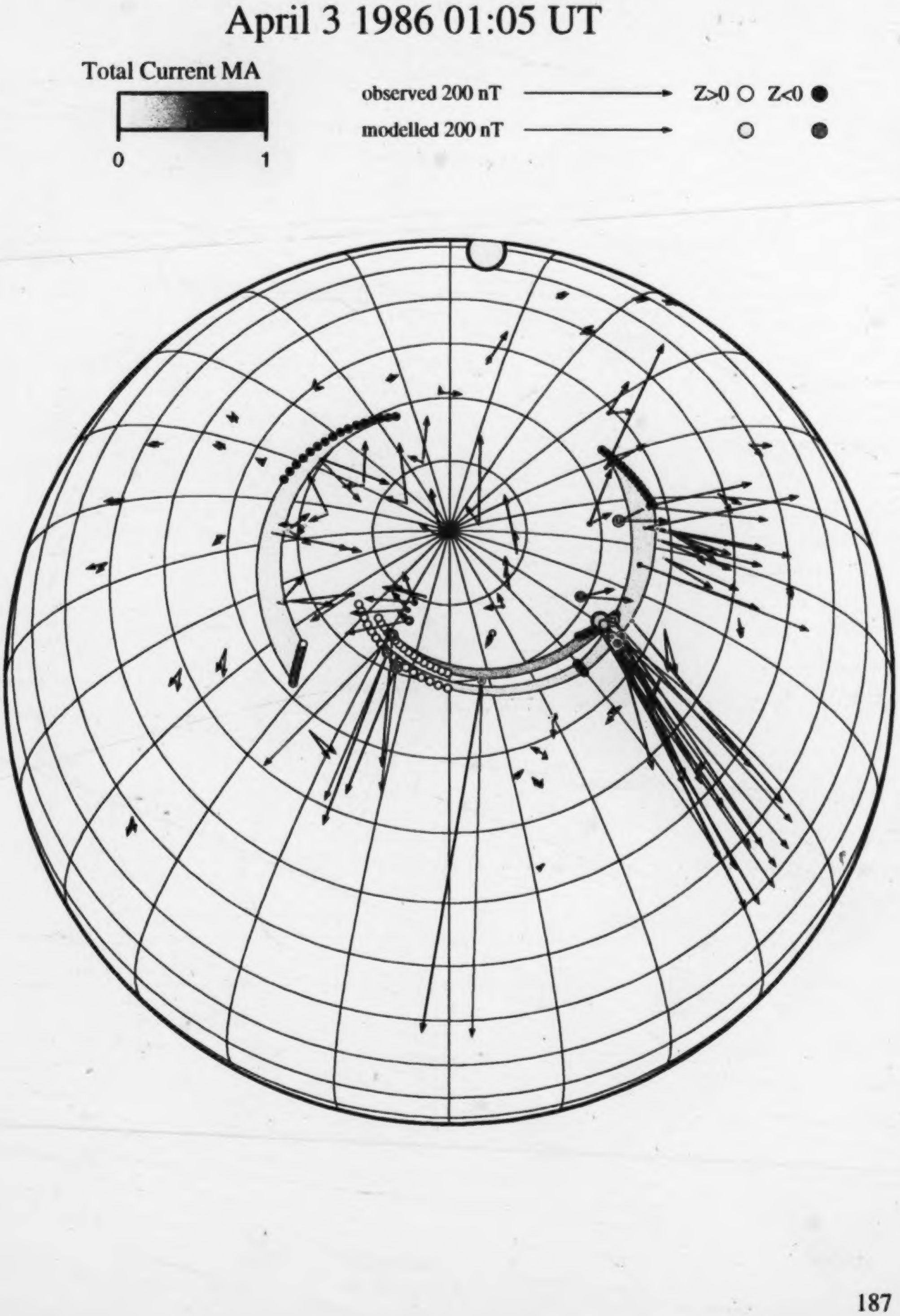

April 3 1986 01:05 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●
○ ●

Figure 6.31 Modelling results at 0110 UT April 3 1986.

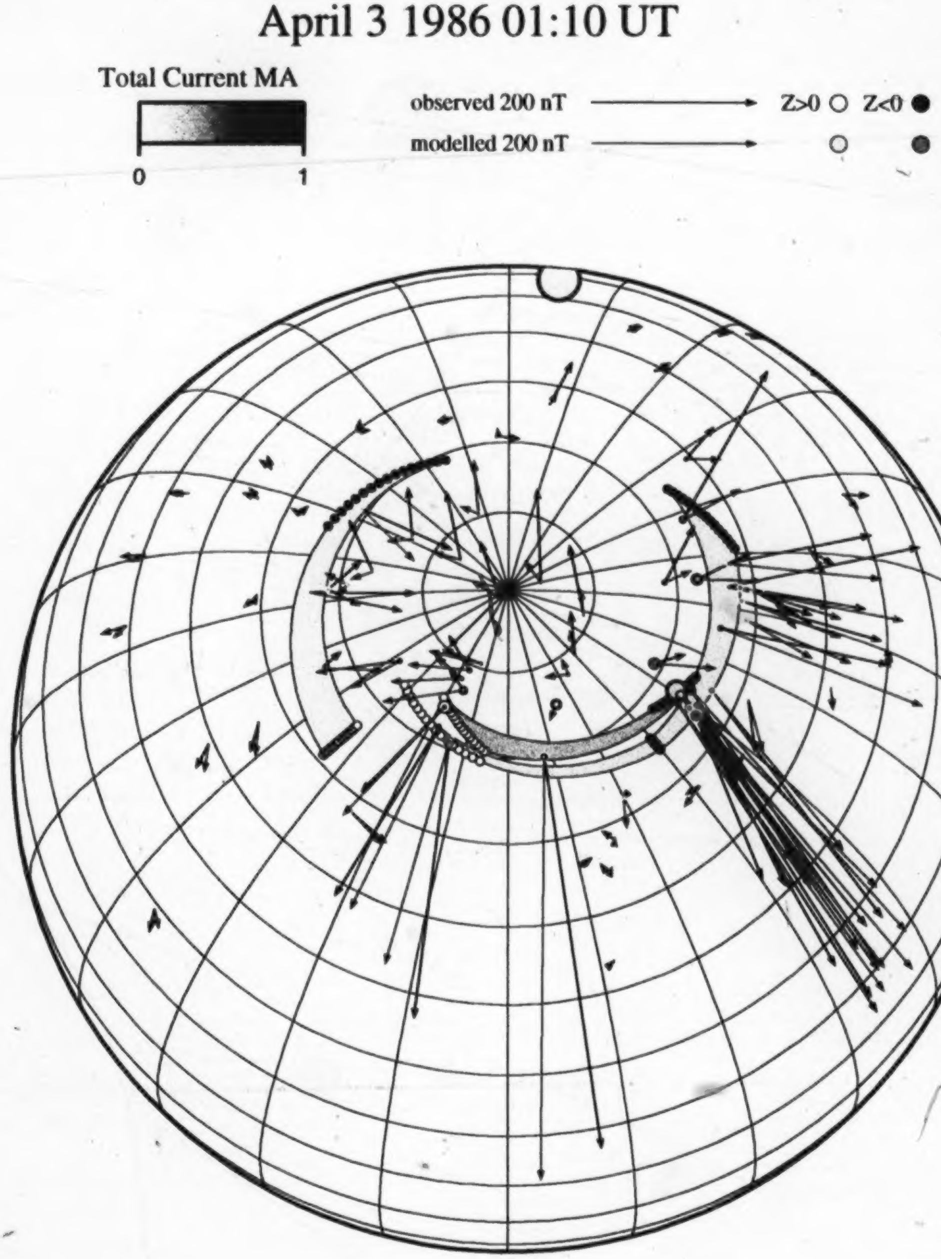

April 3 1986 01:10 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.32 Modelling results at 0115 UT April 3 1986.

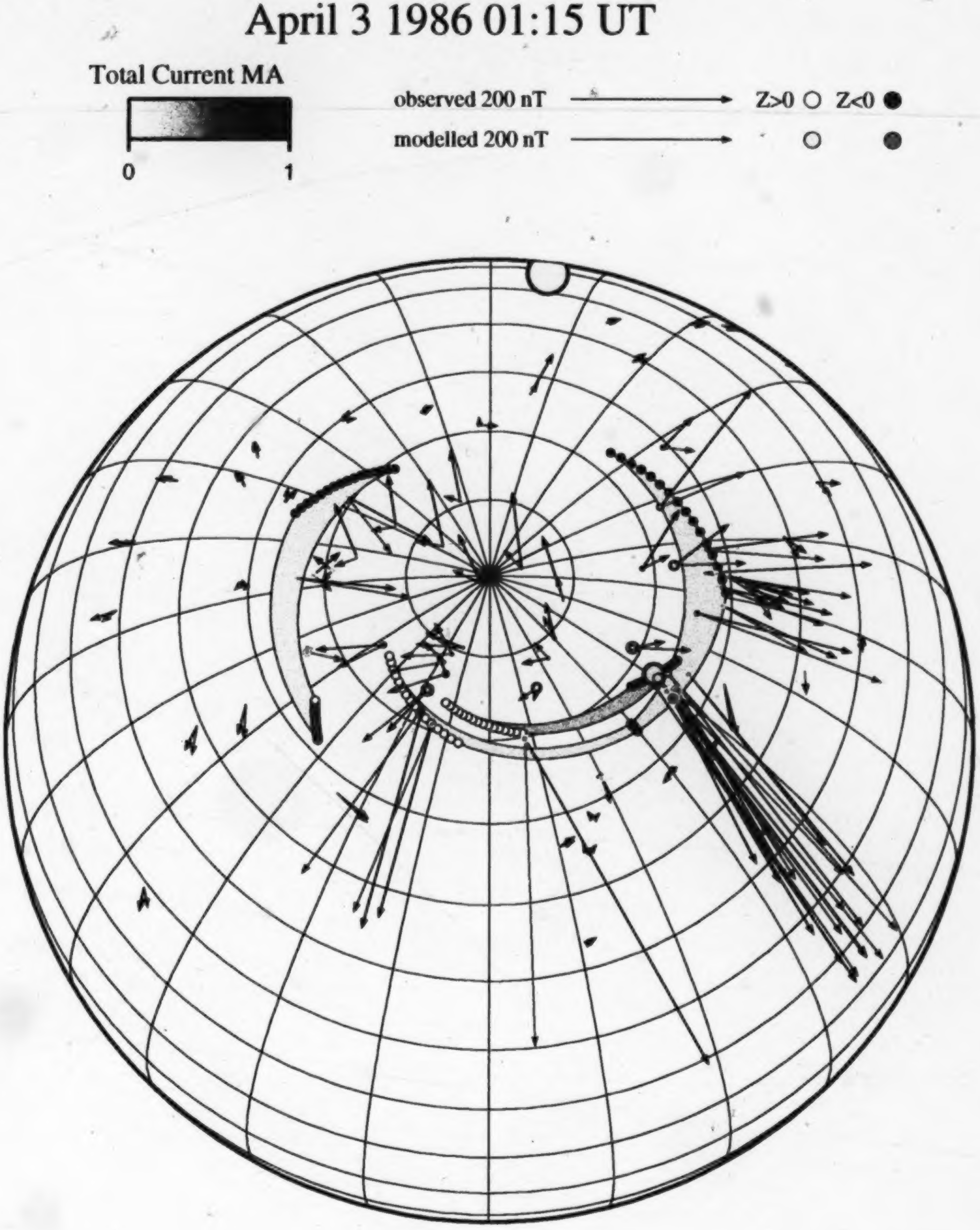

April 3 1986 01:15 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.33 Modelling results at 0120 UT April 3 1986.

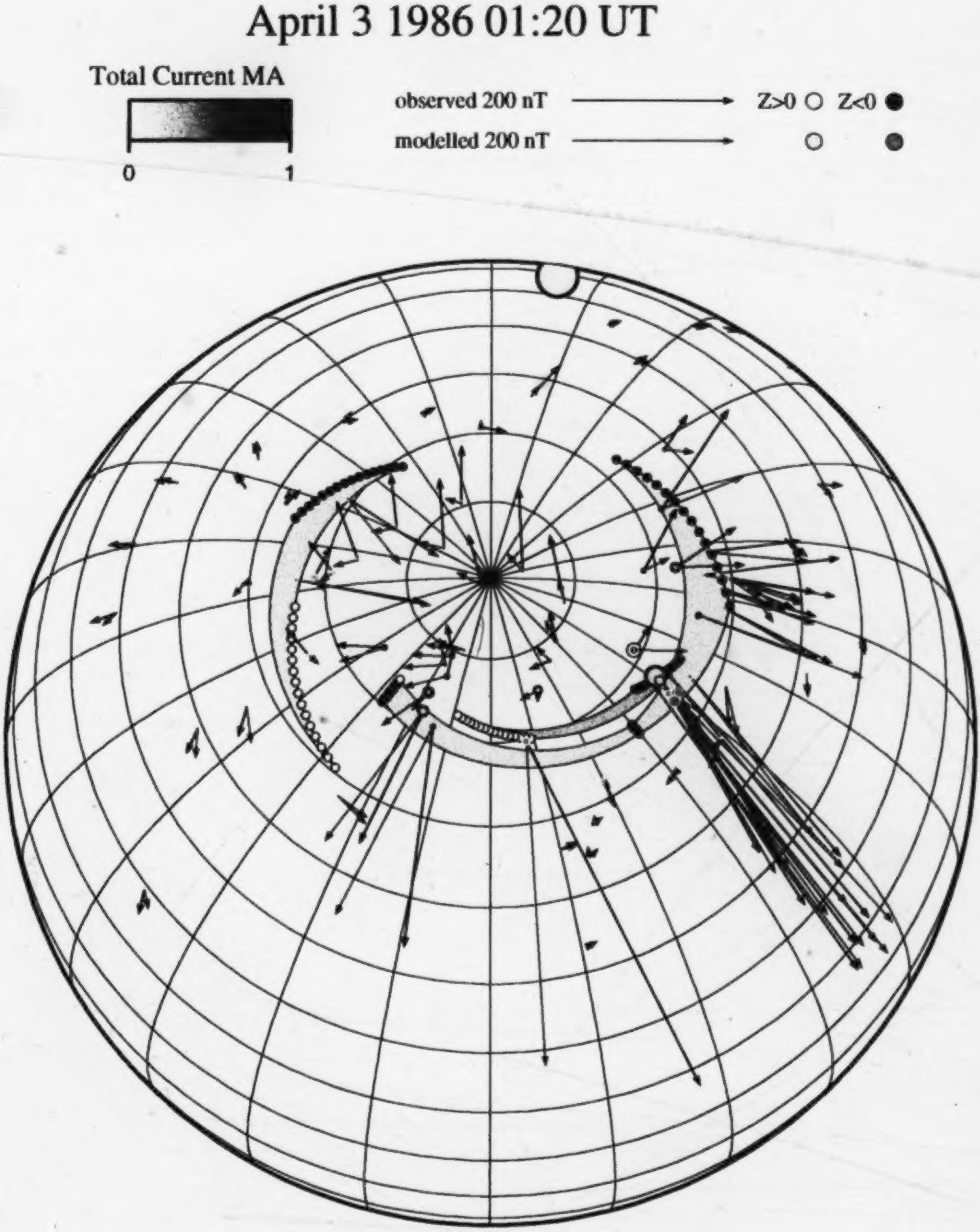
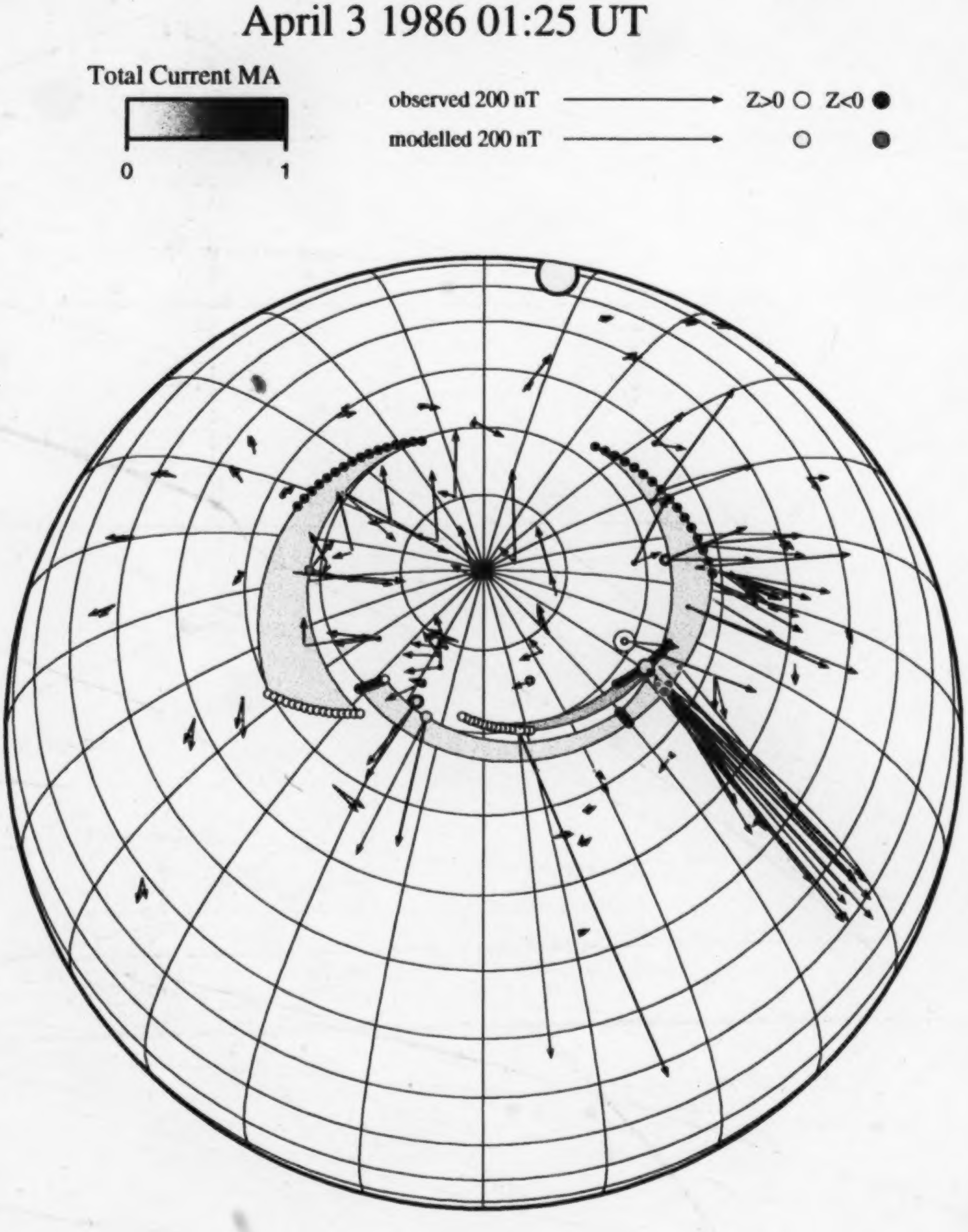


Figure 6.34 Modelling results at 0125 UT April 3 1986.


April 3 1986 01:25 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

In Figure 6.28, as in the figures showing the stations and initial conditions, centred dipole coordinates are used and the grid is 10° in latitude and 15° in longitude. The central meridian is at 60° and the latitude of the projection centre (the 'nadir') is at 70°. Downward field-aligned current is indicated by black circles, upward by white circles. The overall intensity of each current system is indicated by shading, from 0 MA (white) to 1 MA or more (black). Perturbations at ground stations are indicated by vector arrows, black for observed perturbation, red for modelled perturbations. The Z component is indicated as a circle at the base of the vectors (thus at the station in question). Z circle size is proportional to Z component but scale is such that the circle radius is one-tenth that which would be associated with an equal amount of horizontal component. Observed Z is black for negative values, white for positive values. Modelled Z is dark pink for negative values, light pink for positive. Model results have been plotted after the observed perturbations were plotted. In particular for Z, observed results may be obscured. The magnetic longitude of the subsolar point is shown by a yellow circle: its latitude is not that of the subsolar point but is such as to place it on the figure in a useful position. These considerations of data display will apply in subsequent chapters of this document.

In terms of results shown, Figure 6.28 indicates a good agreement between the model and the observations at this time, which effectively blends pre-onset and post-onset conditions due to the five minute time averaging of the data. The morning sector westward electrojet is on the right, incipient substorm current wedge between 90° and about 0°, and the evening sector eastward electrojet on the left. The second portion of the westward electrojet has not been well-constrained or needed for this model and has been placed north of the morning sector electrojet. There are indications that the overall westward current flow starts near noon and perhaps extends somewhat more westward than shown. This conclusion comes from noting the southward perturbation at CPS in the noon sector, and that at IQL in the late evening sector. These two stations do not carry enough weight to force the current systems to move in order to better represent the perturbations observed at them.

In Figure 6.29, the results of modelling at 0100 UT April 3 1986 are shown. This is immediately after substorm onset and perturbations in the midnight sector have strengthened somewhat. The perturbations at IQL are well represented here as the edge effect from upward field-aligned currents in the midnight sector, so that it in fact does likely lie somewhat west of the terminus of the westward current flow. Perturbations in western Greenland are well represented by the currents as shown, and both the current wedge and the western leg of the electrojet have been used effectively in the modelling. The results suggest the doubling of the current flow shown to be real. The substorm current wedge is fed by field-aligned current north of Scandinavia and is well modelled at Iceland (largest arrow). Midlatitude perturbations of the wedge are well represented at subauroral stations in Europe and in eastern North America. The evening sector eastward electrojet is weak but well modelled within North America.

The basic layout of currents in Figure 6.30 is much like that in Figure 6.29. One should note that in none of these figures have polar cap perturbations been well modelled. These models include only the currents explicitly shown: there are no Region 1/2 field-aligned currents in this modelling run.

Figure 6.31 shows the results of modelling at 0110 UT. The wedge is now stronger and wider. There is ambiguity about western extremity of westward electrojet, which has been replaced by wedge as primary current channel over North Atlantic, where there are no stations. At Chelyuskin (CCS), near the northern border of the auroral oval, there appears to be stronger current required than is provided by the model. The westward electrojet is well constrained further west and south by numerous Russian stations, and the claim could still be made that there is a weak westward electrojet at subsolar longitudes. The eastward electrojet in the evening sector has strengthened and northward perturbations there have increased.

There have not been very large changes in proceeding to Figure 6.32, which shows the situation at 0115 UT. There is a wedge/westward electrojet ambiguity as exemplified here. The wedge current has been bled off into space in a poorly constrained region and westward electrojet extended over southern Greenland to satisfy constraint by stations there. Looking ahead to the image in Figure 6.39, a region of upward field-aligned current southwest of Greenland, as indicated here, appears correct, whereas the wedge upward FAC between Iceland and Greenland does not seem to be supported as there is not much luminosity in that region, and luminosity would be expected if upward field-aligned current was present there. Note that the orientation and foreshortening effects differ between the modelling results and the images.

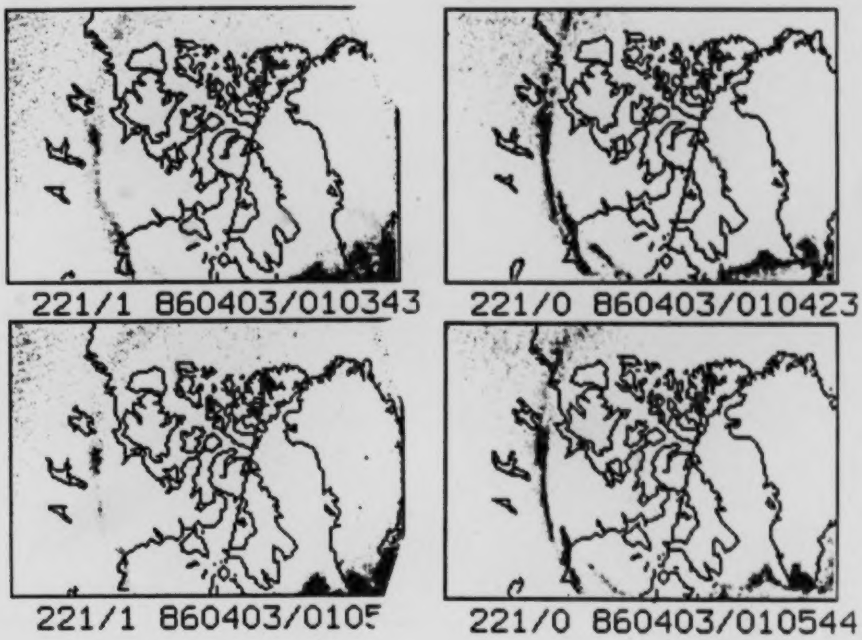
Comparing Figure 6.33 to Figure 6.29, changes in the evening sector electrojet are apparent. It has both strengthened and moved southward by 0120 UT. Stations in Northern Greenland now exhibit perturbations like those of the polar cap, which are not adequately modelled here, representing a northward veering of the perturbations seen at those stations.

By 0125 UT, as shown in Figure 6.34, the western terminus of the 'wedge' at the northern border of the westward electrojet appears to have retreated eastward. Unlike the situation in Figure 6.34, the reasons for this are now obvious, in the form of reduced perturbations in southern Greenland. Polar cap perturbations continue to be poorly modelled, and those in northern Greenland generally share their sunward orientation. The least well represented stations near the electrojets are those on the poleward border of the westward electrojet, and notably BJN (Bjørnøya), which has, since 0120, greatly strengthened its positive Z and southward horizontal perturbations. These are the sign of a strong current south of the station. The fact that the electrojet as placed models very well the perturbations south of BJN suggests that this strong current is not, in fact, the main electrojet current.

Figures 6.35 through 6.42 show inverted greyscale Viking images for the time period 0104 to 0126. Comparison to model results reveals that the placement of the evening sector upward field aligned currents by the modelling program corresponds closely to regions of emission in the western part of the images. The upward field aligned currents are likely associated with precipitating electrons (primaries being in the 3 to 7 keV range [Vallance Jones *et al.*, 1987]), which would cause emissions and thus this identification seems sensible. There is a weak eastward electrojet configuration in the evening sector, as revealed by modelling. What is seen in the images as a wavy but sharp band of emission is likely for the most part associated with Region 1 upward currents at the northern border of the auroral oval, while the more southerly Region 2 downward currents, due to upwelling cold electrons, do not produce significant emission. The Viking field of view does not encompass much of the region of upward currents in the morning sector - what can be seen shows an active and broad auroral oval. The substorm current wedge in the modelling run described above was characterized by being north of the driven system electrojets with a longitudinal extent of roughly 90°, from about 120° geomagnetic to 30° geomagnetic. Throughout much of the modelling run the upward field-aligned currents associated with it and with the westward electrojet were found to have about 5° latitudinal extent and to be located between about 15° and 60° geomagnetic longitude. The precipitation regions seen by Viking have extent of about 5° in latitude and appear to move with time from 40° geomagnetic to about 0° geomagnetic. The vignetting-corrected Viking images allow a stepwise westward motion of the wedge-associated upward field-aligned current to be seen. A step from about 40° to about 20° occurs at starting at about 0107 UT (Figure 6.36), with emission having shifted largely to northern Labrador (about 20°) by about 0109 (Figure 6.37). A second step to 0° occurs at 0116 (Figure 6.39). The stepwise westward development is not present in the modelling results, but the region where they occur is not well covered by magnetic stations. The sequence shown here ends with a general brightening of evening sector auroras and development shortly thereafter of a complex but short-lived structure in the 0° geomagnetic region (see Liu and Rostoker [1993]). In the period 0116 to 0118 all magnetic components at Poste-de-la-Baleine (-10° centred dipole longitude) decrease abruptly: the Z component by 150 nT (Figure 6.43). These signatures are those of a 'wedgelet' developing north of the station, but the abrupt drop in H component associated with an expansive phase onset is never seen and the perturbations are short-lived. This indicates that the wedge never did expand over this station, but that somewhat north of it there was an active element. The midlatitude signatures of the substorm current wedge also end at about this time, as may be verified by inspection of Figure 6.23. The images are now described in more detail.

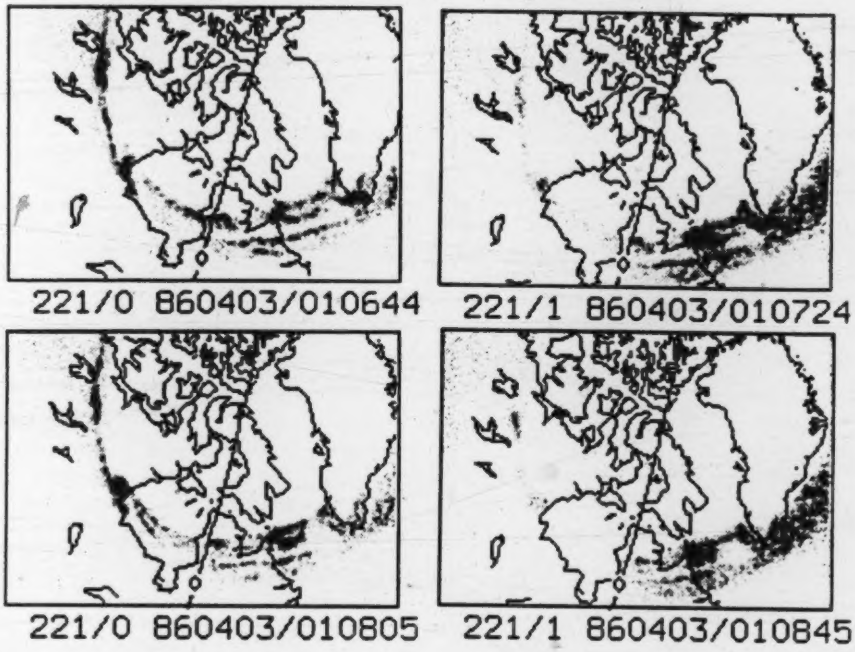
Figure 6.35 shows Viking images from both cameras. As in subsequent images, the LBH camera is indicated by a '221/0' at lower left, the 1304 camera by a '221/1'. Upward field-aligned current associated with the evening sector eastward electrojet dominates in the western region (left) and corresponds closely to that shown in modelling. Bright spots in southern Greenland (lower right) are associated with the substorm current wedge already active at this time. In this and subsequent images vignetting in the Viking camera has been corrected to enhance features on the left and right edges; the 0° EDFL geomagnetic meridian has been superposed for comparison with model results. The modelling results are presented in roughly the same orientation as are the images. Comparison is also aided by noting that Greenland has a western coastal magnetometer chain at about 30° longitude (see inset of Figure 6.25).

Figure 6.35 Viking images from 0104 to 0106 UT April 3 1986.



In Figure 6.36, a bright area has formed at the northern tip of Labrador. That area brightens in this sequence while luminosity over Greenland fades. It is also clear that the auroral oval is bifurcated, which is consistent with results in the earlier stages of modelling. In Figure 6.37, when the wedge is strongest (ca. 0110) as indicated by modelling and by midlatitude perturbations, luminosity over Labrador continues to increase and to spread east-west. Brightening of more southerly region off the Labrador coast starts to form a pronounced double structure. In Figure 6.38, double structure off Labrador dominates. The distribution of luminosity is consistent with the modelling result that most current flows over southernmost Greenland. The double structure's northern emission region is not fed by currents crossing Greenland north of the westward electrojet. However, there is also no indication of enhanced emission to correspond to upward field-aligned current between Iceland and Greenland, as noted above. Notable in this as in other images in this section is the brightness and structure seen in the evening sector in Camera 0.

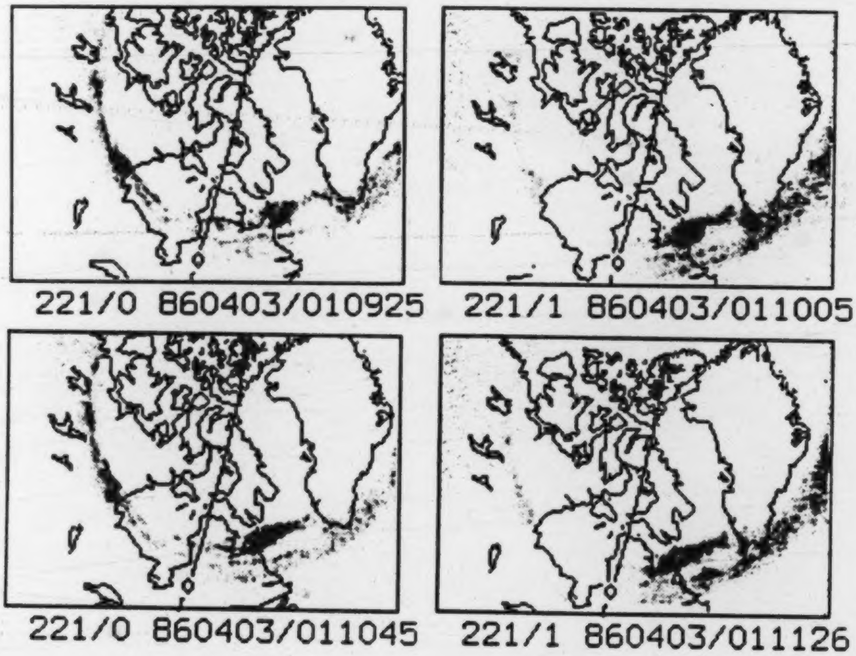
Figure 6.36 Viking images from 0107 to 0109 UT April 3 1986.



221/0 860403/010644 221/1 860403/010724

221/0 860403/010805 221/1 860403/010845

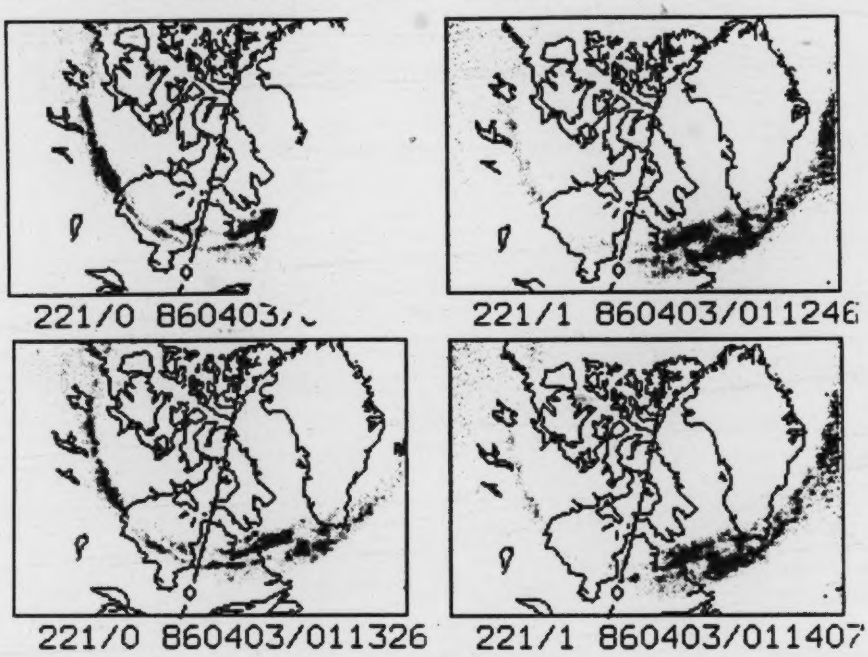
Figure 6.37 Viking images from 0109:25 to 0111:26 UT April 3 1986.



221/0 860403/010925 221/1 860403/011005

221/0 860403/011045 221/1 860403/011126

Figure 6.38 Viking images from 0112 to 0114 UT April 3 1986.



A further westward jump of the main particle precipitation region is seen in Figure 6.39. Activity shifted to the region over Poste-de-la-Baleine at about 0° geomagnetic. This corresponds to the eastward limit of the evening sector eastward electrojet as determined magnetically. The brightest auroral activity also jumped southward and Camera 0 emissions in the entire evening sector strengthened. The declining stage of substorm current wedge activity is shown in Figure 6.40, for the time just prior to 0120 UT. Camera 0 evening sector emissions remained strong and the emission about 0° geomagnetic took on the high Camera 0/Camera 1 ratio to be discussed below. A strong perturbation seen at Poste-de-la-Baleine at this time has a large negative Z component, negative X_M component, and a change in sign of Y_M from a positive to a negative perturbation, as seen in Figure 6.43. The final images available while the wedge was still active are shown in Figure 6.41. Morphologically the 0° geomagnetic region is characterized by double structure with north-south aligned emission regions. The sequence terminates with fading emission in the south. By the end of this sequence the magnetic signature of the substorm current wedge is gone. In the following image, emission patterns in the evening sector (left) have returned to those of 20 minutes earlier (see Figure 6.35) while those in the Greenland region are featureless and faint.

Figure 6.39 Viking images from 0115 to 0117 UT April 3 1986.

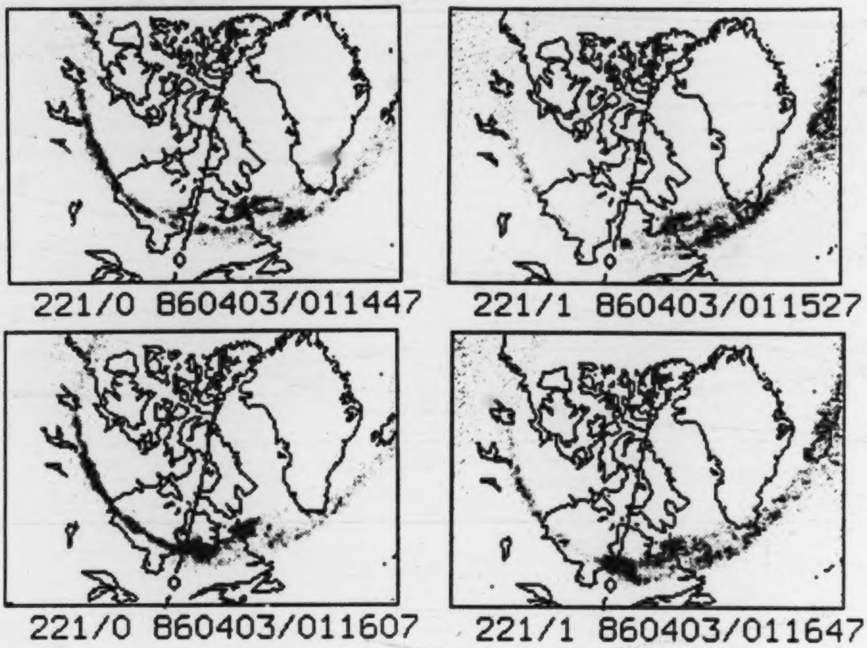

221/0 860403/011447
221/1 860403/011527
221/0 860403/011607
221/1 860403/011647

Figure 6.40 Viking images from 0117:28 to 0119:28 UT April 3 1986.

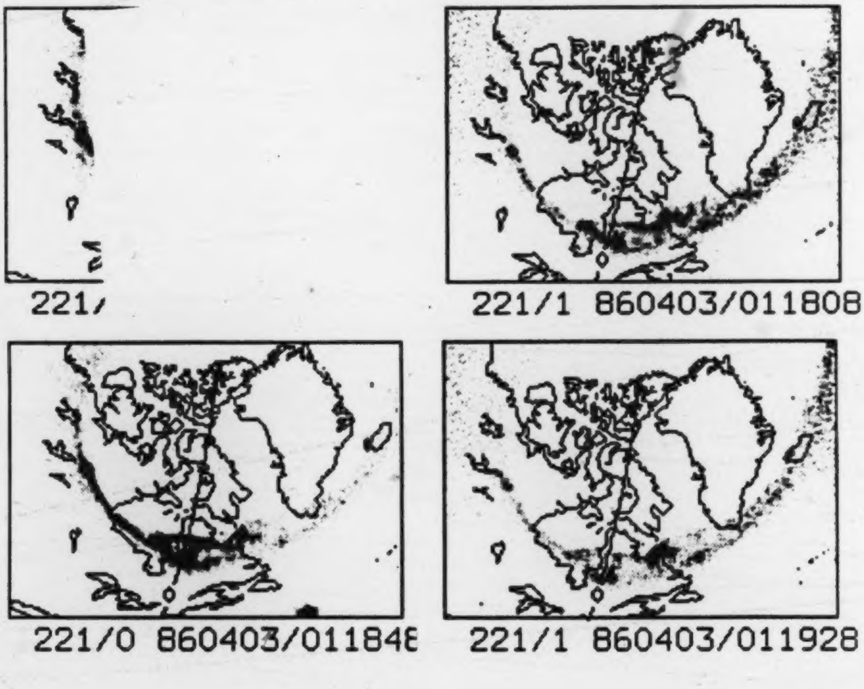
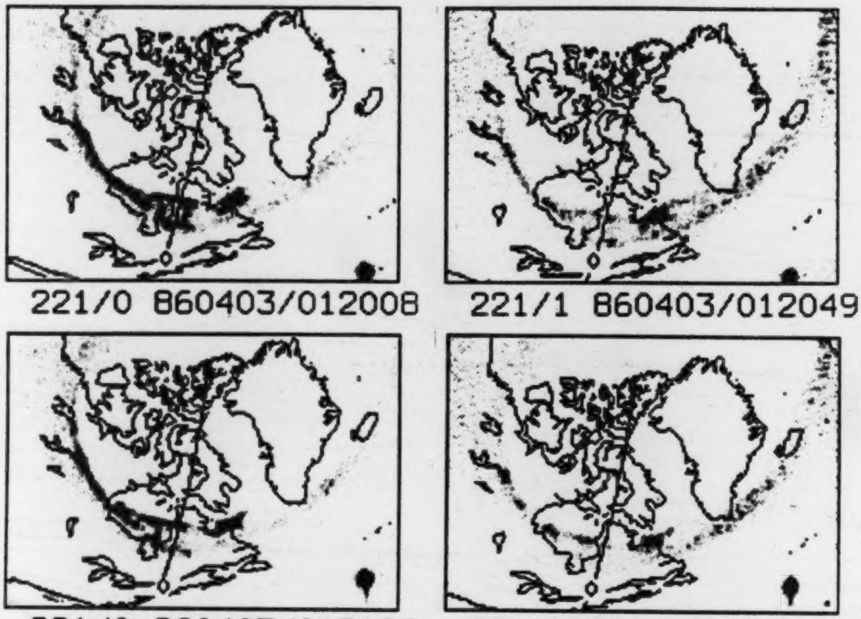

221/
221/1 860403/011808
221/0 860403/011848
221/1 860403/011928

Figure 6.41 Viking images from 0120 to 0122 UT April 3 1986.



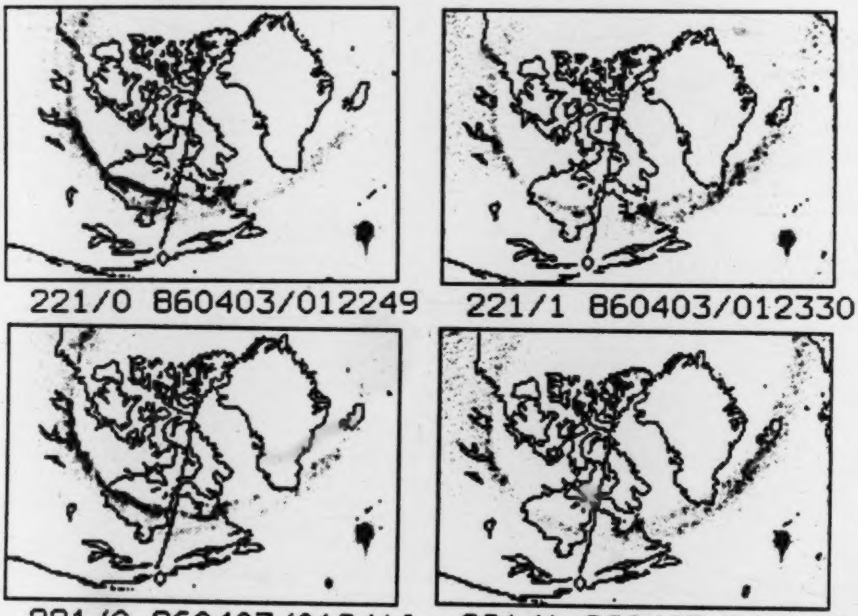
221/0 860403/012008

221/1 860403/012049

221/0 860403/012129

221/1 860403/012209

Figure 6.42 Viking images from 0123 to 0125 UT April 3 1986.



221/0 860403/012249

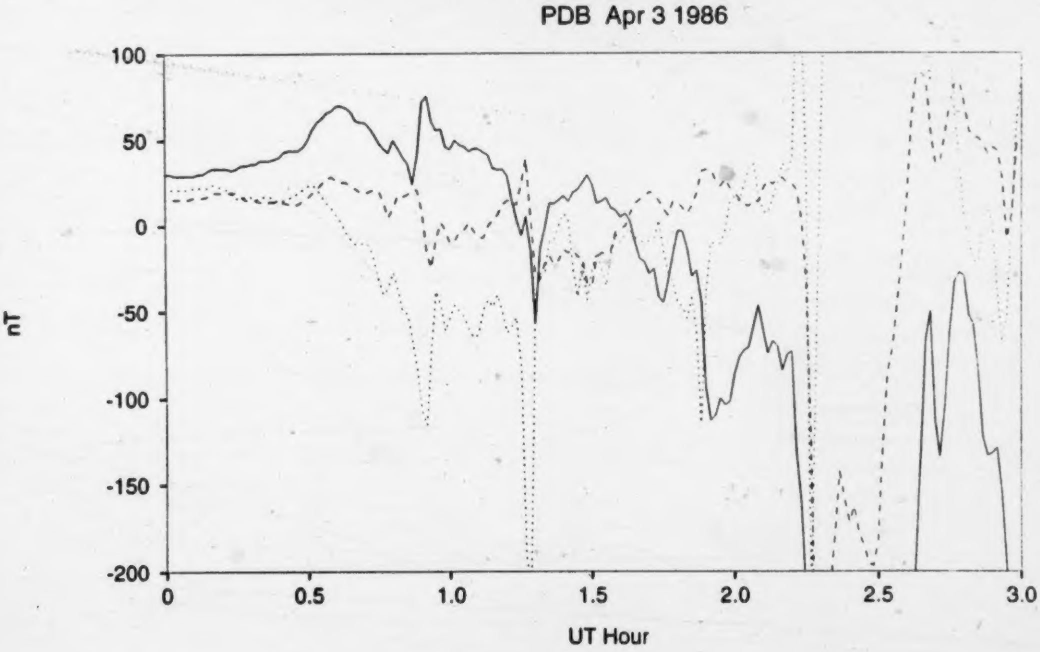
221/1 860403/012330

221/0 860403/012410

221/1 860403/012450

It is useful to now re-examine Figures 6.35 through 6.42 in conjunction with modelling results suggesting with which current system the zones of emission are associated. Throughout, it may be seen, the zones of emission clearly associated with the evening sector eastward electrojet and corresponding Region 1 currents have a low ratio of Camera 1 to Camera 0 luminosity. This effect was first noted by Anger *et al.* [1987a]. Recalling (Section 3.c.1) that Camera 0 detected mainly emissions from the LBH bands of molecular nitrogen, while Camera 1 detected neutral atomic oxygen (OI) at 130.4 nm (1304 Å) wavelength, a low Camera 1/Camera 0 ratio may be referred to as a low 1304/LBH ratio. The differences at these wavelengths could be due either to absorption above the emitting layers or to differences in the incident exciting particle spectrum [Vallance Jones *et al.*, 1987]. For the CDAW 9A event (see Chapter 7), Persson *et al.* [1992] found that the western region of the substorm current wedge (the westward travelling surge region) was characterized by low 1304/LBH ratios and high Hall/Pedersen conductivity ratios. Both of these are attributed to precipitation of 3 to 4 keV electrons. High 1304/LBH ratios are found when the upward field-aligned current associated with the wedge is mainly over Greenland (Figures 6.36 and 6.37) and when it is primarily over or near northern Labrador (Figures 6.38 to 6.40). A high ratio of 1304/LBH is typical of the morning sector [Persson *et al.*, 1992]. However, after the jump to the Poste-de-la-Baleine region (Figure 6.40 and subsequently), the spectral characteristics in the western part of the current wedge changed to be similar to those of evening sector forms. Shortly thereafter the substorm current wedge ceased to exist.

Figure 6.43 Perturbations at Poste-de-la-Baleine from 0 to 3 UT on April 3 1986. Components shown are X_M (solid), Y_M (dashed) and Z (dotted).



The midlatitude signatures (Figure 6.23) suggest that the wedge currents were present until about 0130, but westward expansion in the images did not appear to occur after about 0120. Although the modelling results did not show any stepwise expansion of field-aligned current regions, the wedge was found to retreat after 0120, and to be in decline after that. The events described are suggestive of a close connection between the substorm current wedge and the electrojets of the driven system. Stepwise motion of the emission regions associated with the wedge appeared to stop when it approached MLT midnight, spectral changes occurred, and the evening sector emissions strengthened, along with an increase in the total current in the evening sector. Unfortunately, poor station placement does not allow this interaction to be modelled in more detail with the automated forward modelling technique.

Table 6.9 Stations used in making Russian meridian profiles.

Code	Name	Geodetic Latitude	Geod. Long.	CD Latitude	CD Longitude	CD Dec.
HIS	Heiss Island	80.6	58.0	71.5	156.5	-28.2
ZHE	Mys Zhelaniya	77.0	68.6	67.4	157.7	-19.0
BEY	Ostrov Belyj	73.3	70.0	63.7	155.8	-16.0
KHR	IZMIRAN KHR	71.1	66.8	61.9	152.5	-16.0
TAM	Tambei	71.5	71.8	61.8	156.0	-14.4
YSH	IZMIRAN YSH	69.8	60.8	61.2	147.6	-17.4
SYH	IZMIRAN SYH	70.1	72.5	60.4	155.7	-13.5
CKA	IZMIRAN CKA	68.5	73.6	58.8	155.8	-12.5
YSL	Yar-Sale	66.8	70.8	57.4	153.1	-12.8

Another useful forward modelling approach will now be discussed. Latitude (meridian) profiles can be made in various meridians for which data are complete: here these comprise the Churchill meridian in North America, the 120° magnetic meridian which includes the EISCAT stations, and the Russian chain at about 155° magnetic whose stations are listed in Table 6.9. The Russian chain shows interesting behaviour at onset time which may be well studied with the simplest of forward models: that of a long westward electrojet. Kisabeth [1972], in his introductory sections, shows the latitude profiles associated with westward electrojets. These illustrations are not repeated here: the reader is referred to Section 5.b.1 for a discussion of the perturbations from similar systems. Kisabeth used systems of east-west extent 20°, whereas those used here are longer (120°) and more appropriate to a driven system electrojet. By comparison to global modelling, which is fully three-dimensional, the use of latitude profiles is essentially two-dimensional. His use of systems which are too short does not invalidate Kisabeth's work, but here it has proven practicable to use systems of appropriate length. The placement of the systems in longitude is not critical: the aim is simply to place the systems appropriately so that essentially two-dimensional modelling may be done. The constraint rules described in Chapter 5 were used to specify a system having purely east-west ionospheric current flow in a regions whose latitudinal borders were variable. Such a model electrojet is incapable of producing Y_M component perturbations at its midpoint, although small Y_M perturbations due to the remote up and down field-aligned currents may be seen if the meridian chain modelled is not exactly in the electrojet centre. A long system places these field-aligned currents far away from a modelling region near its centre and thus effects modelled are almost entirely those of ionospheric currents.

Figure 6.44 Modelling results for electrojet borders of Russian chain on April 3 1986.

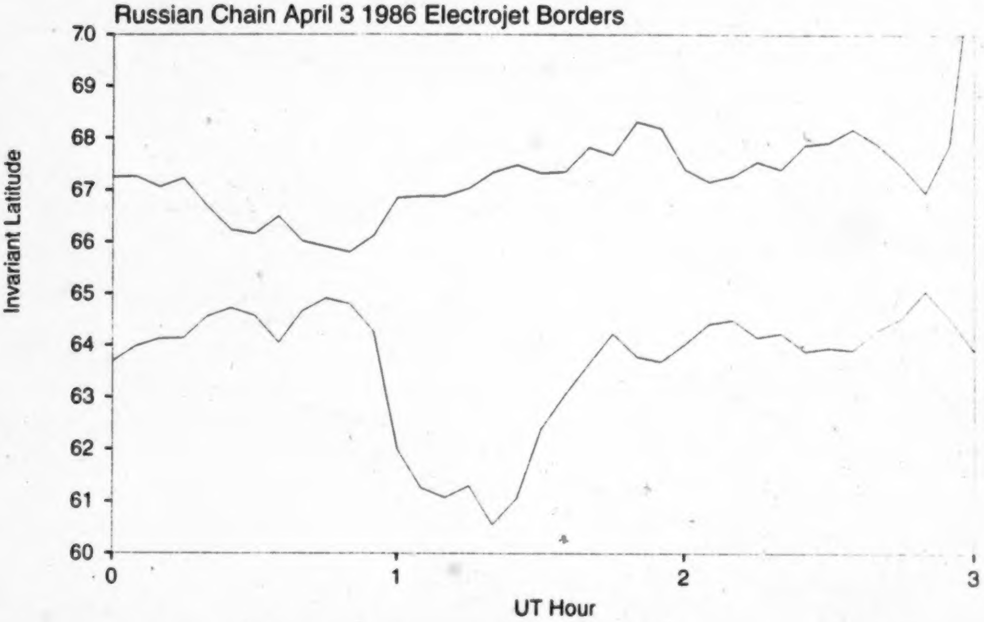


Figure 6.44 shows the electrojet borders calculated in the Russian meridian from a simple electrojet model with downward current at 180° and upward current at 60°. Stations used in making these profiles, listed in Table 6.9, span the centred dipole latitudes from about 57° to 72°. The northern border remains relatively constant at about 66° to 67° during the onset period. The seeming 'necking together' of the northern and southern borders from about 0020 to 0055 will be discussed below. The most dramatic feature of this latitude profile solution is the southward motion of the southern border of the electrojet at onset time, i.e. at 0055 UT. This is followed by a northward retreat of the southern border at approximately 0130 UT. No further large motions are observed at a subsequent major onset (i.e. at about 0215 UT). The last point indicates northward motion of the northern border but may represent breakdown of the forward model. As will be seen below, the northernmost point, Heiss Island, is generally not well represented by the simple electrojet and has relatively strong perturbations (see for example Figure 6.45). This can cause problems for the convergence of the routine. Given the simplicity of the forward model, it is useful to examine the individual latitude profiles to verify the solutions and to assess how much confidence and meaning to attach to them.

Figure 6.45 Latitude profile of Russian chain at 0 UT April 3 1986. Data points are shown with circles for X_M, upward triangles for Y_M, and downward triangles for Z. Model fit is shown with lines: solid for X_M, dashed for Y_M, and dotted for Z.

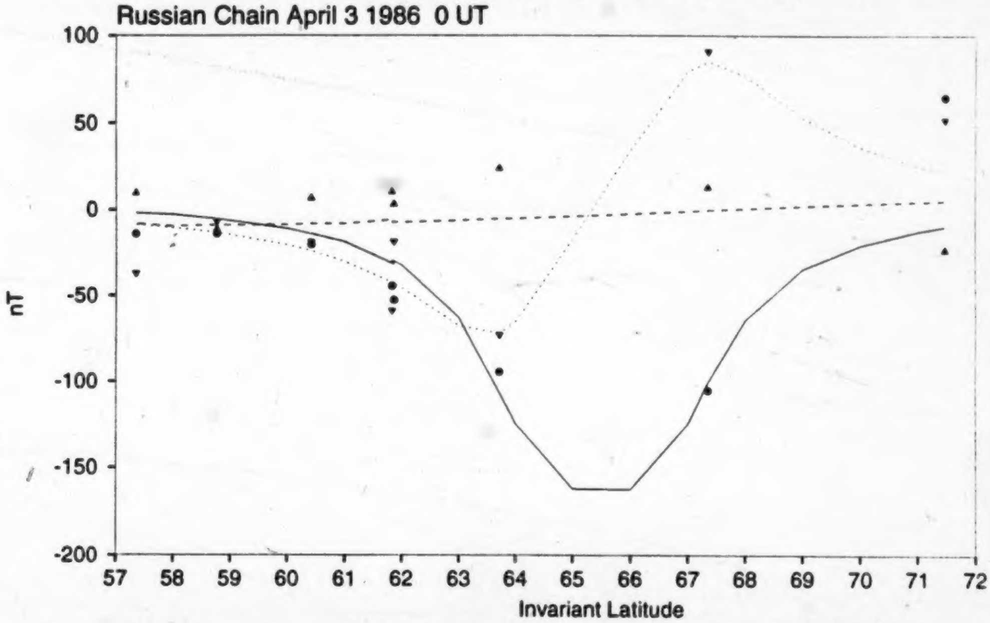


Figure 6.45 shows the profile of the Russian chain at 0 UT, which is early in the growth phase leading to the 0055 UT onset. In this and subsequent diagrams, observations are indicated at stations as circles for X_M, upward triangles for Y_M, and downward triangles for Z. Model fits are also plotted, both at stations and at uniform 1° intervals, and are joined by lines. Following the general convention used above in this work, these lines are solid for X_M, dashed for Y_M, and dotted for Z. Since the model represents perturbations at all points in space, there are more points plotted for the model than for observations, and the model may contain extrema which were not observed at stations. For example, in Figure 6.45 there is a minimum in X_M near 65° where there are no stations. The reality of this minimum is attested to by the good fit in X_M and in Z where there are stations available. The Y_M component as modelled in this and subsequent profiles is that of distant field-aligned currents only: the electrojet model, at constant but variable latitude, cannot reproduce local Y_M components. The X_M and Z components at all stations except the most northerly (Heiss Island) are well reproduced by the model. It is quite likely that Heiss Island is in the polar cap at this time and thus under the influence of currents which are not in the forward model. Its contribution to the weighted χ^2 has been small (as that of any single outlier data point should be) and thus the model has converged on a solution valid at the other stations. That the model does represent a good fit for the electrojet is supported by noting the good fit at the Z extrema and that these extrema correspond to the electrojet borders as indicated at this time in Figure 6.44. The use of the Z extrema to roughly indicate electrojet edges is well known [Kisabeth, 1972; Kisabeth and Rostoker, 1974].

Figure 6.46 Latitude profile of Russian chain at 0050 UT April 3 1986. Fit indicated corresponds to 'necked down' electrojet.

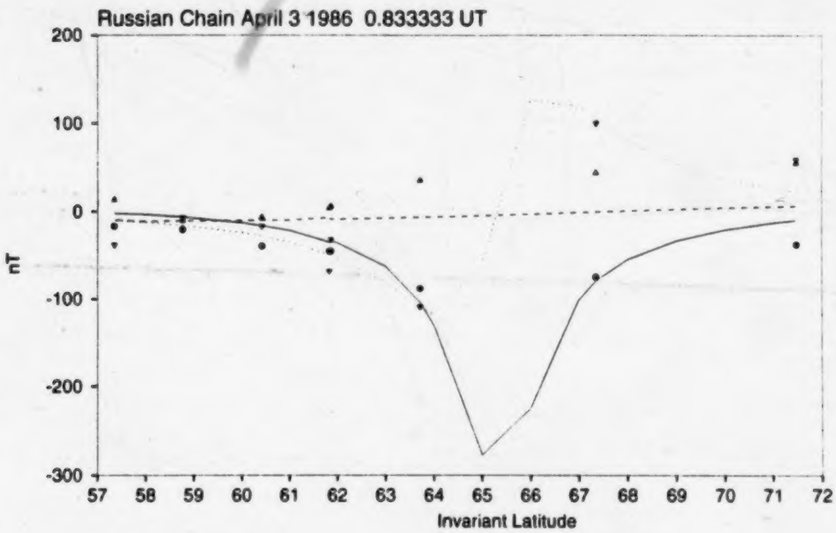


Figure 6.46 is the profile at 0050 UT, when the electrojet borders are apparently 'necked together'. Careful inspection shows that most data points are similar to their counterparts at 0 UT with the exception of those from Heiss Island, which now conform to the perturbations expected from only the electrojet. The slightly larger Z magnitude to the south (the northern Z is larger in a uniform electrojet), and the relative change in the lowest X_M points, suggest that the solution is in fact correct and that 'necking' did occur. However, the effects are subtle and it is possible that the changes at Heiss Island have had an effect. By contrast, the southward motion of the southern electrojet border at onset time is surely significant, as may be seen by inspection of Figure 6.47. Here the X_M component would suggest that the electrojet or parts of it were found further south, while the Z component is not as well modelled. This reflects a possibly more complicated structure of the electrojet. The distribution observed corresponds well to one with an intense northern border and a diffuse distribution of current further south as modelled by Kisabeth [1972, p. 65]. Another possibility investigated by Kisabeth [1972, p. 33] is that of a distant current masquerading as an electrojet. That possibility can be dismissed here since the large Y_M component perturbations (approaching those of X_M) which he found in such a case are not seen here. Further, the three-dimensional modelling (see Figures 6.32 to 6.34) also shows electrojet southward motion despite explicit inclusion of the wedge system. The southward excursion of the southern border was short-lived, approximately corresponding to the duration of the wedge associated with the 0055 onset.

Southward motion in the morning sector was noted in DE auroral images by Craven and Frank [1987], but seemed not to be as sharp as in the case here. Kamide and Kokobun [1996] indicate that poleward motion of the auroral oval (and in particular of the northern border) characterize substorms triggered by northward turning of the IMF, but that if the IMF remains southward after onset (as is the case here), then equatorward motion may characterize the dusk and dawn sectors. In this case, some equatorward motion had already occurred before onset in the dusk sector, and this larger equatorward shift after onset happens in the dawn sector.

Figure 6.47 Latitude profile of Russian chain at 0105 UT April 3 1986. Rapid southward motion of the southern boundary is obvious in comparing to other figures.

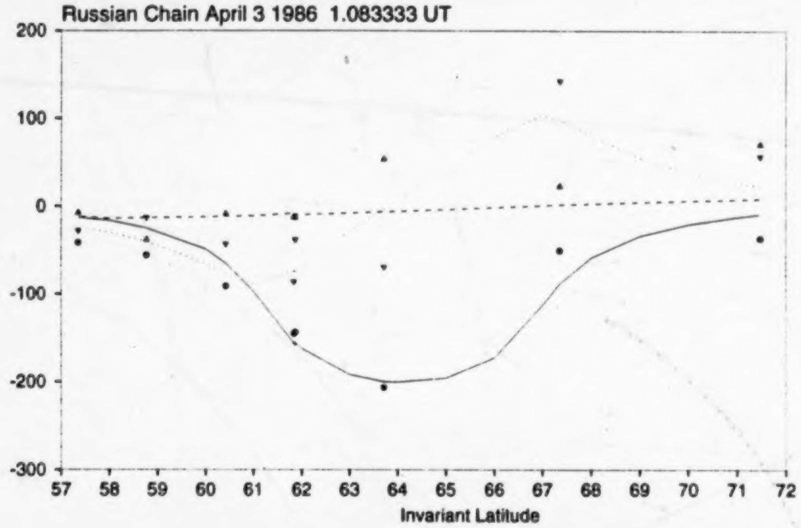
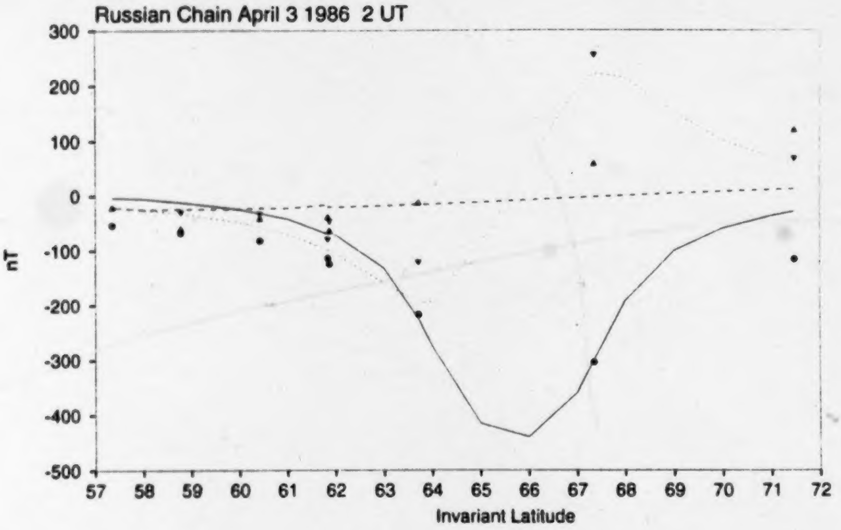


Figure 6.48 shows the chain at 0200, well after the decline of the 0055 wedge (although others may have occurred as seen below). The form of this profile strongly resembles that at 0 UT except that Heiss Island no longer seems poorly represented. Another small difference is that there appears to be weak distributed current at the southern border. The border solution as seen in Figure 6.44 is basically identical to that at 0 UT. The magnitude of the current has increased as is evident by the much larger perturbations at 2 UT as compared to 0 UT.

Figure 6.48 Latitude profile of Russian chain at 2 UT April 3 1986. Electrojet position is comparable to that at 0 UT, but intensity has approximately doubled.



The total current is the third variable used in latitude profile fitting (the borders are the other two) and is shown in Figure 6.49: it is seen to have approximately doubled at this meridian between 0 UT and 2 UT. The actual increase in total current across the meridian while the 0055 UT substorm current wedge was active does not seem to have been as great as that after its disappearance. In the figure, the rise in current at 3 UT is related to the spurious width increase already noted: if the width increases without much change in perturbation at most of the stations, then the total current must increase in proportion (approximately) to the width. It can thus be concluded that the rise in current at 0300 UT is spurious.

Further points may be useful concerning operation of the fitting routine. Here, the Z values at approximately 67° are very important: if they are neglected, there is considerable ambiguity in the latitude profiles. A prior run was done in which this ambiguity caused poor convergence, essentially due to a poor choice of scale factor parameter. The χ^2 values for this earlier, rejected run, are shown in Figure 6.50 and are seen in all cases to be larger than those of the fits which were accepted. The effect of single stations may also be seen in the poor representation of Heiss Island in the profiles. Not only early on in the sequence, but also at times towards 3 UT, the representation of the perturbations at this station is not good. A spurious northward 'leap' of the poleward electrojet border appears to occur near 3 UT. Inspection of the magnetograms and latitude profiles (not all of which are shown) shows that Heiss Island generally registers perturbations during this event which are more extreme (positive before onset and negative after onset) than those due to an otherwise well-fitted uniform electrojet. In addition, the Z component at this station is generally subdued. These could both be indications of anomalous behaviour due to near-surface conductivity. It may thus be appropriate to weigh this station less in solutions, as has been done in AMIE runs [B. Emery, private communication 1995]. In this set of runs this has not been done, which appears to have introduced problems only at 3 UT. Anomalous behaviour at this station is more pronounced during the CDAW 9A event (Chapter 7) and the weighting was thus reduced in that case. Another possibility suggested by these data is that there is an enhanced current at the northern border of the electrojet. That would help to explain the tendency for northern border stations to have larger than expected perturbations as seen in the global modelling.

Figure 6.49 Current across Russian chain at 0050 UT April 3 1986.

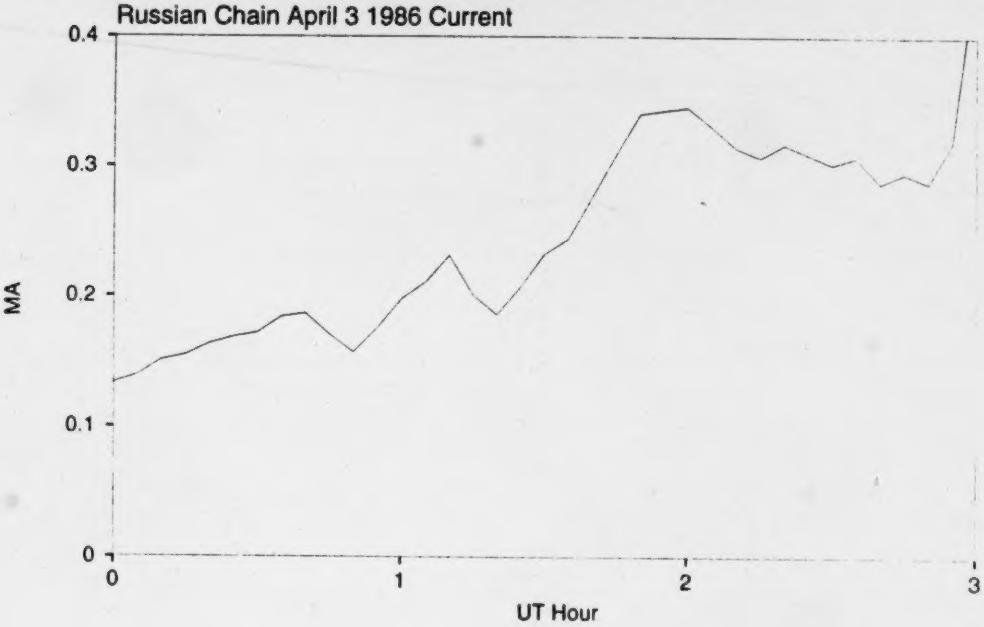
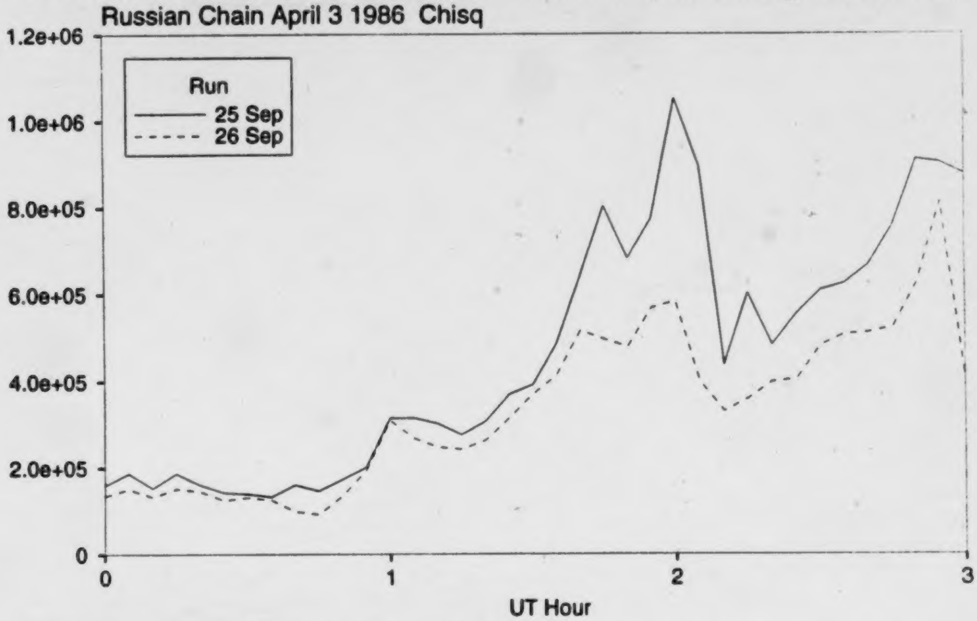


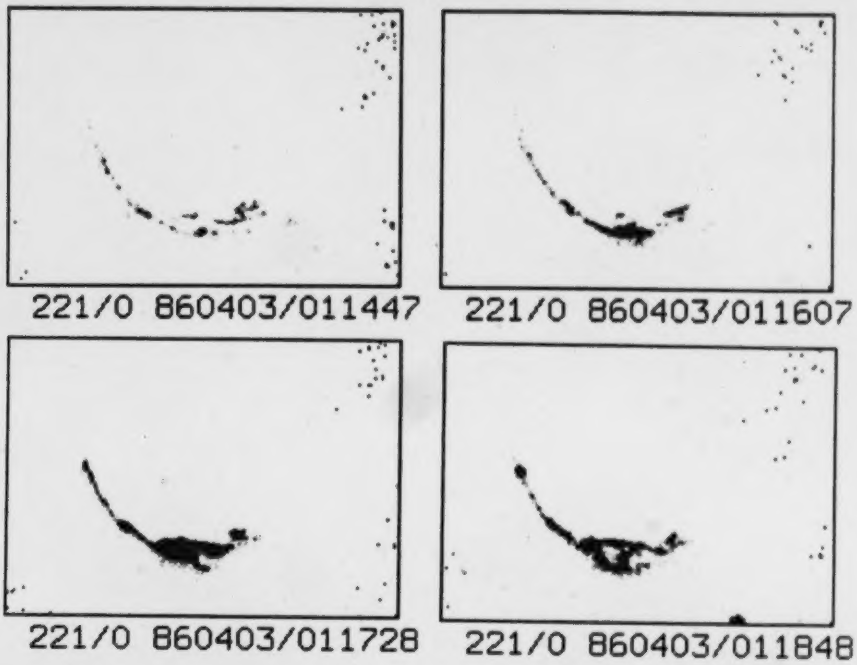
Figure 6.50 χ^2 for two fitting runs to Russian chain. Bottom line (dashed) corresponds to results shown in Figures 6.44 through 6.49. Solid line has larger χ^2 in all cases and corresponds to a rejected run in which incomplete convergence occurred.



e. Small-scale behaviour in western wedge

Mention has been made several times of a 'wedgelet' detected primarily through perturbations at Poste-de-la-Baleine. The signature of the 'wedgelet' consists of a Z deflection of about -140 nT, an X_M deflection of about -70 nT delayed by about one minute, and a Y_M deflection of about +40 nT, as may be verified by inspection of Figure 6.43. These perturbations peaked shortly before 0118 UT (1.3 UT) and the entire sequence is shown in Figure 6.51 as imaged by Viking. These images repeat some shown in Figures 6.39 and 6.40 but with a greyscale more appropriate for studying the brightest features. For clarity, no continental outlines are shown. Reference to those figures will show that the brightenings were localized in the midnight sector (for clarity Figure 6.51 also does not show coordinate lines). The region of interest shows complex structures immediately before onset, and the auroral oval contains at least two bright arcs trending east-west. It is the more equatorward of these which has a kink in the location at which major brightening later occurs (see Figure 6.38 in which this kink is immediately east of the midnight meridian). This arc system merges continuously with the evening sector arcs.

Figure 6.51 Viking images of 'wedgelet' from 0116 to 0118 UT April 3 1986



The Automated Forward Modelling routine may be used to attempt to find the source of the perturbations and relate them to the images. Perturbations distinct from wedge signatures are of a useful size only at Fort Churchill and Poste-de-la-Baleine (Figure 6.43). Small perturbations observed at Ottawa are not used here. Even at Fort Churchill the only distinct signature is an approximately 30 nT negative excursion of X_M. The question then becomes one of determining a current system which can give the observed perturbations using only two stations. Obviously, the details of the eastern part of a wedge-like system cannot be determined from these stations, which are located toward the west of the wedge. Thus, the longitude of the eastern downward current was chosen and held fixed (at 60°) and the latitudes of this current sheet were constrained to be those of the upward current. A modelling run was done with the position of only the upward current system from the 'wedgelet' to be determined along with the current in the system. This leaves 5 parameters to be determined from 6 observations, when Y_M and Z perturbations at Fort Churchill are taken to be zero. The result is that the current flow is between 67.8° and 68.8°, with upward field-aligned current leaving the ionosphere at these latitudes between -15° and -18.5°. The system is centred about 2° north of Poste-de-la-Baleine, with upward current about 7° west of the station. The total current in the system is .23 MA. In this best-fit model, the Z component at Poste-de-la-Baleine is well matched at about -125 nT, while the magnitudes of the X_M components at both stations are overestimated by about 50%. The Y_M component at Poste-de-la-Baleine is well matched by about 44 nT of perturbation, and the near-zero Y_M and Z perturbations at Fort Churchill were modelled at about 6 nT each. In the images, the centre of the brightest patch at

0117:28 was measured to be 3° north of Poste-de-la-Baleine. Activity of emission regions in the images suggests, much as do the perturbations, that there was a very brief expansion of a region of westward current flow to near and slightly west of Poste-de-la-Baleine. Overall, the simple model also supports the view that a stepwise expansion of the wedge into the Hudson's Bay region briefly took place. These results must be discussed in context with the global modelling results. It will be recalled (see Figures 6.32 and 6.33) that the global models indicated that the substorm current wedge did not extend this far west. In fact, it was not possible for it to cross Greenland, and only the morning sector westward electrojet intruded to near Post-de-la-Baleine (although north of the station). Once again an ambiguity is suggested between contributions from the substorm current wedge and the westward electrojet.

f. Northerly onset

Modelling for the period 0130 to 0200 UT is shown in Figures 6.52 to 6.58, based on initial conditions which included a northerly 'wedge' as shown in Figure 6.26. These initial conditions are identical to those used in modelling the 0055 UT onset, and this sequence is the continuation of that modelling run. The 0130 to 0200 UT period is typified by an active northern border of the morning sector westward electrojet as modelling results indicate by keeping the 'wedge' in that location while giving a good fit. Nevertheless, the solutions do not correspond very closely to the initial condition: a large change has been made by the modelling routine to obtain these solutions. Even in other modelling runs in which a different initial condition was used, results were not vastly different from those shown. The presence of a gap between westward and eastward electrojets is a reproducible feature of the modelling.

Figure 6.52 shows an overall similarity to Figure 6.34, the overall perturbations not having changed much in the period of five minutes separating them. The most noteworthy change is difficult to see in the figure due to many data arrows overlapping: the horizontal perturbation at BJN has strengthened and become more directly southward. This continuing tendency has caused a redistribution of current in the model for five minutes later, as shown in Figure 6.53: the 'wedge' current system has moved northward to cover the station BJN with a dramatic improvement in the match of the Z and horizontal components. The latter is difficult to see, but the modelling file shows that the observed (X_M,Y_M,Z) perturbations at BJN at 0130 were (-173,-27,295) versus modelled values of (-65,24,84), representing 19% of the total weighted error in the model. These respective component sets were (-194,-67,248) and (-196,8,177) at 0135, so that the station accounted for only 4.6% of total weighted error. The routine responded to the perturbations at BJN once they became large enough, making a significant change in current systems and drawing attention to activity at the station which otherwise might have gone unnoticed. The strengthened current near this station remains in Figure 6.54, which illustrates the situation at 0140. The major change in going to this model is in the evening sector and is discussed on the next text page.

Figure 6.52 Modelling results at 0130 UT April 3 1986.

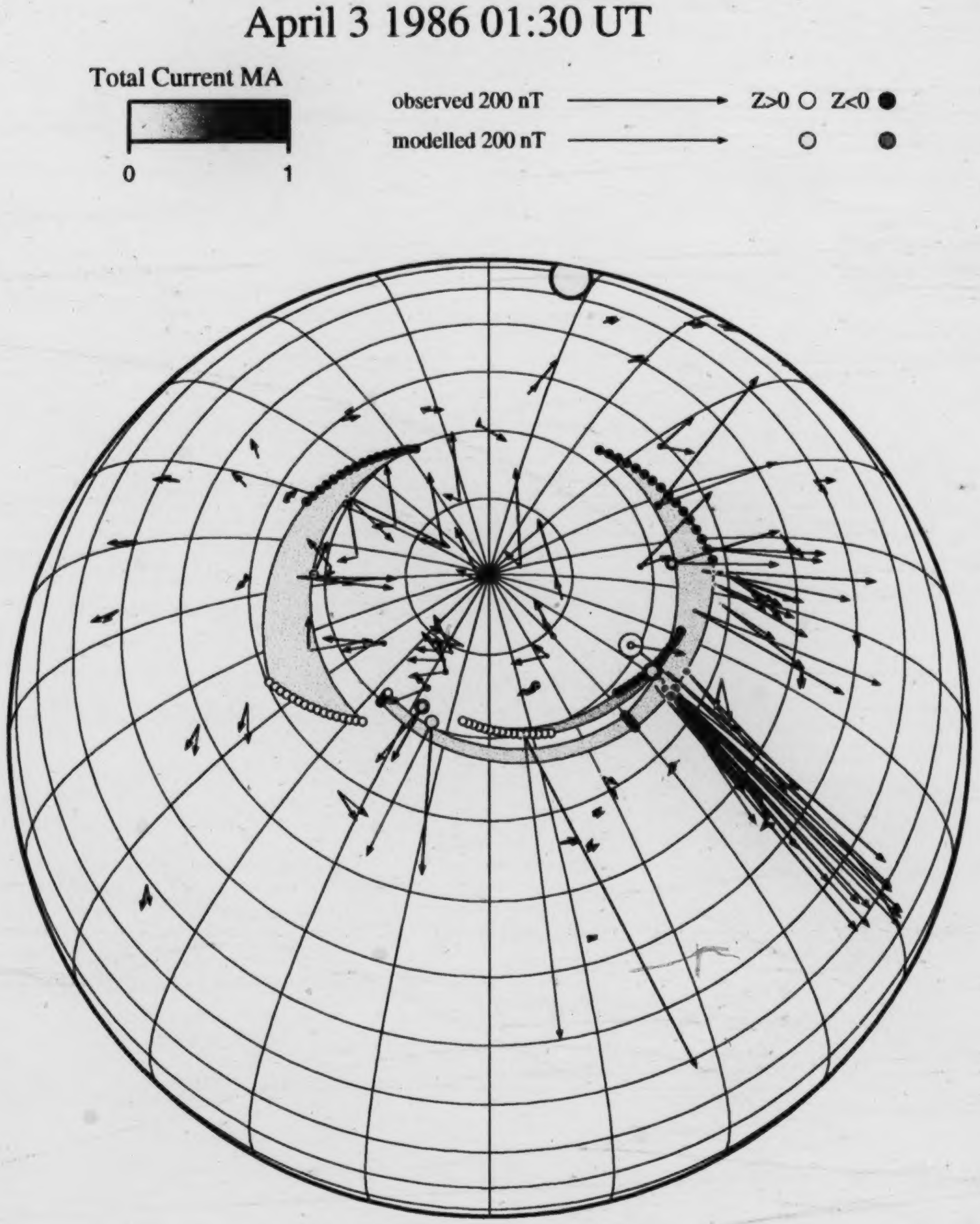

April 3 1986 01:30 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0
Z<0

Figure 6.53 Modelling results at 0135 UT April 3 1986.

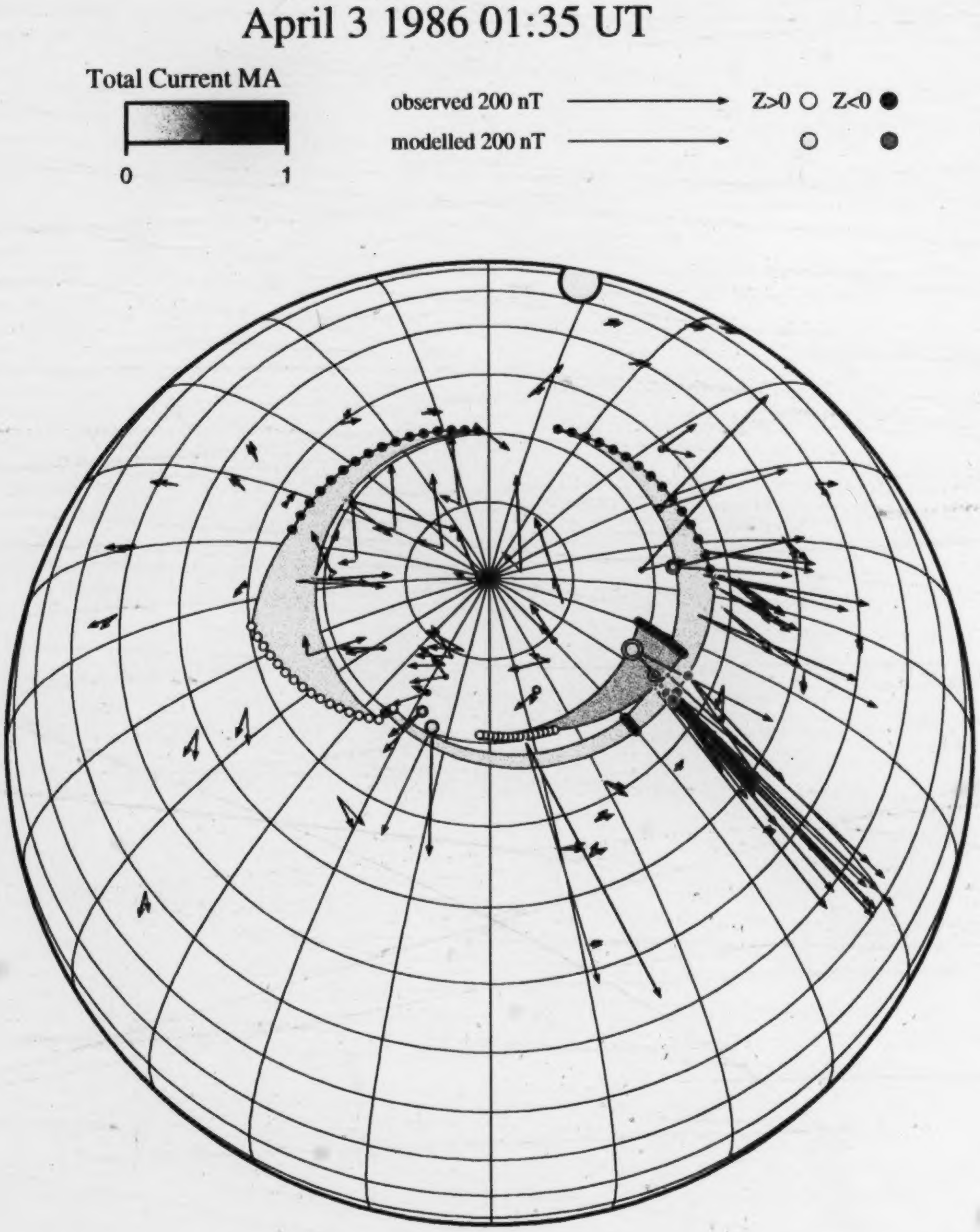

April 3 1986 01:35 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●
○
●

Figure 6.54 Modelling results at 0140 UT April 3 1986.

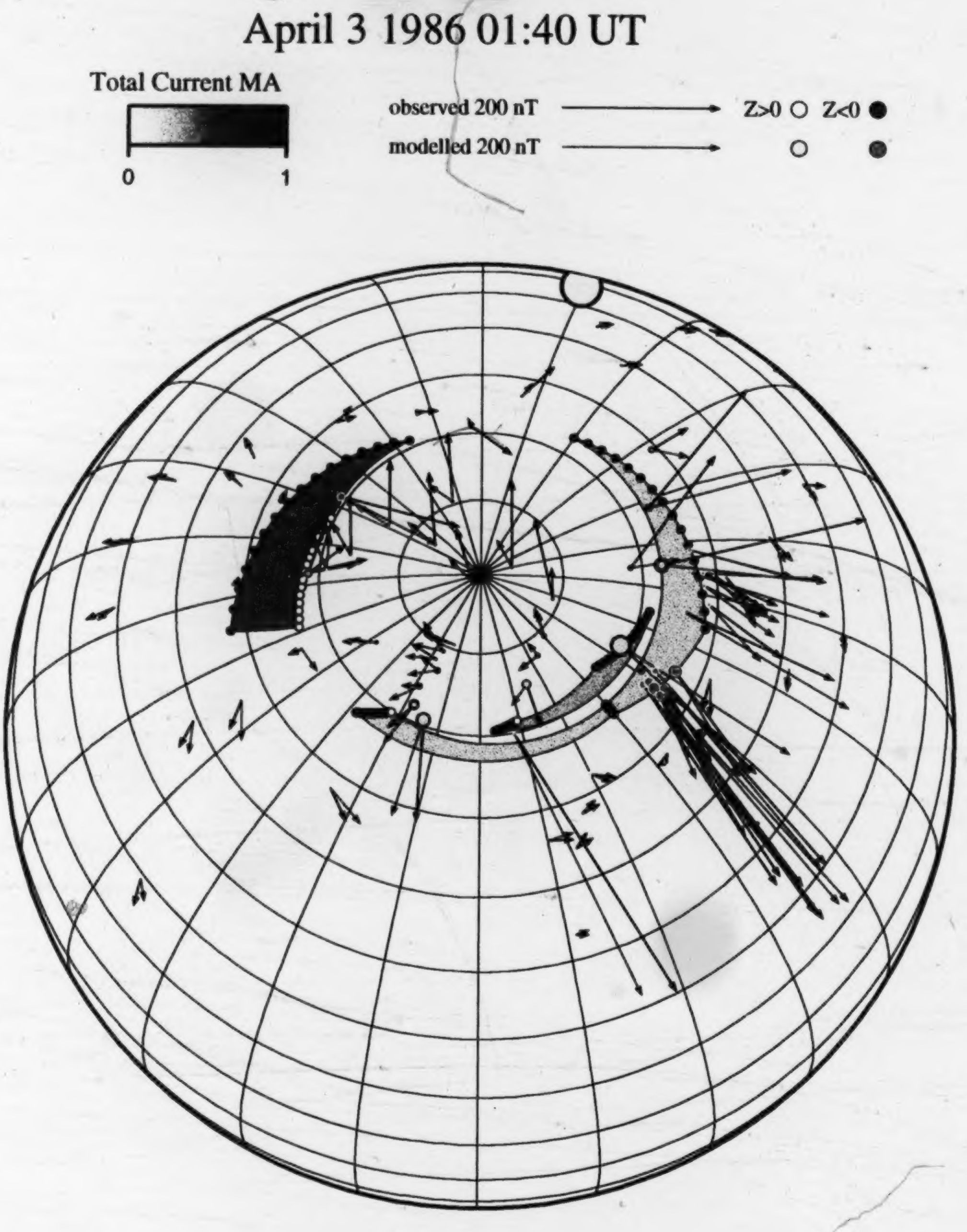

April 3 1986 01:40 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.55 Modelling results at 0145 UT April 3 1986.

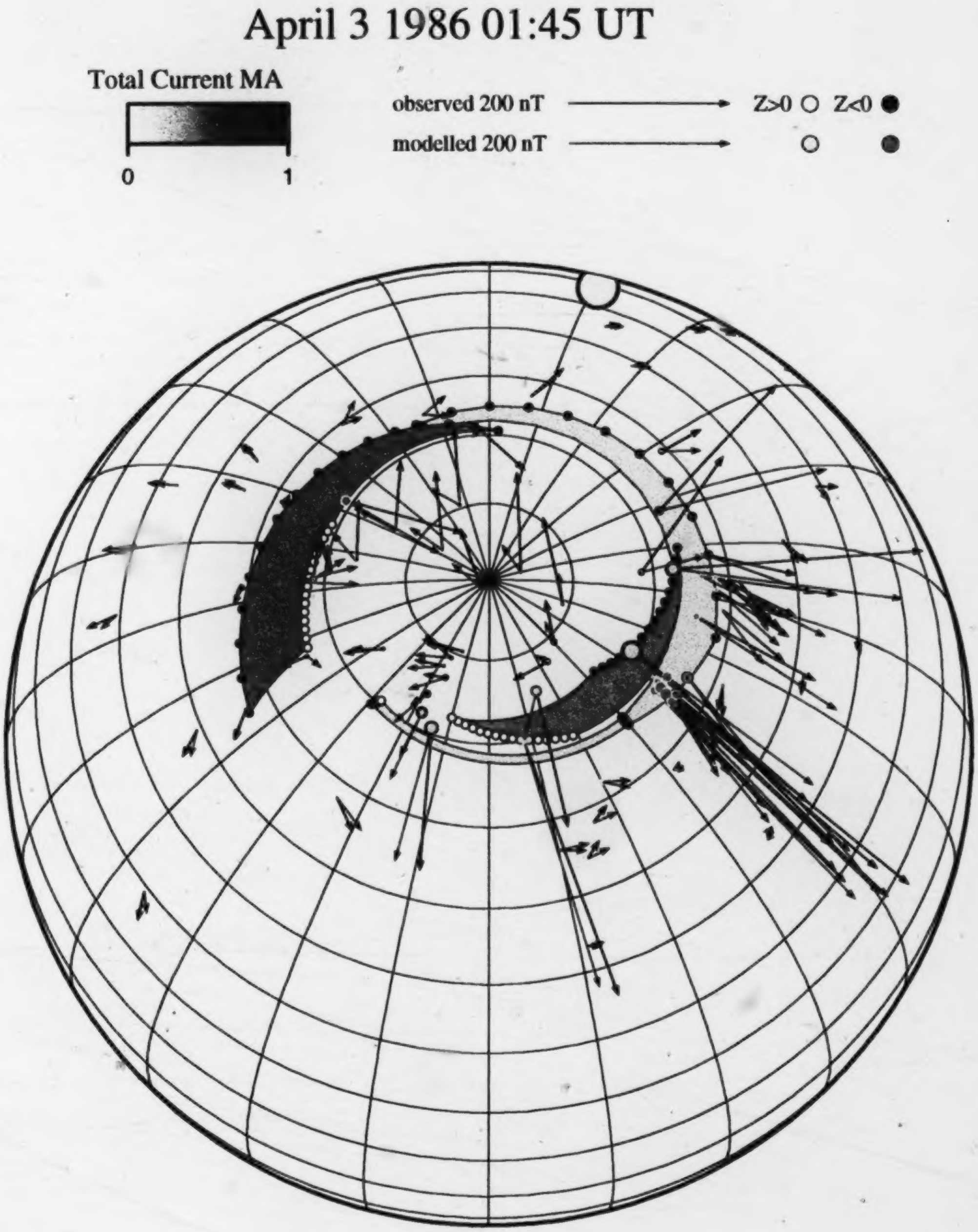

April 3 1986 01:45 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.56 Modelling results at 0150 UT April 3 1986.

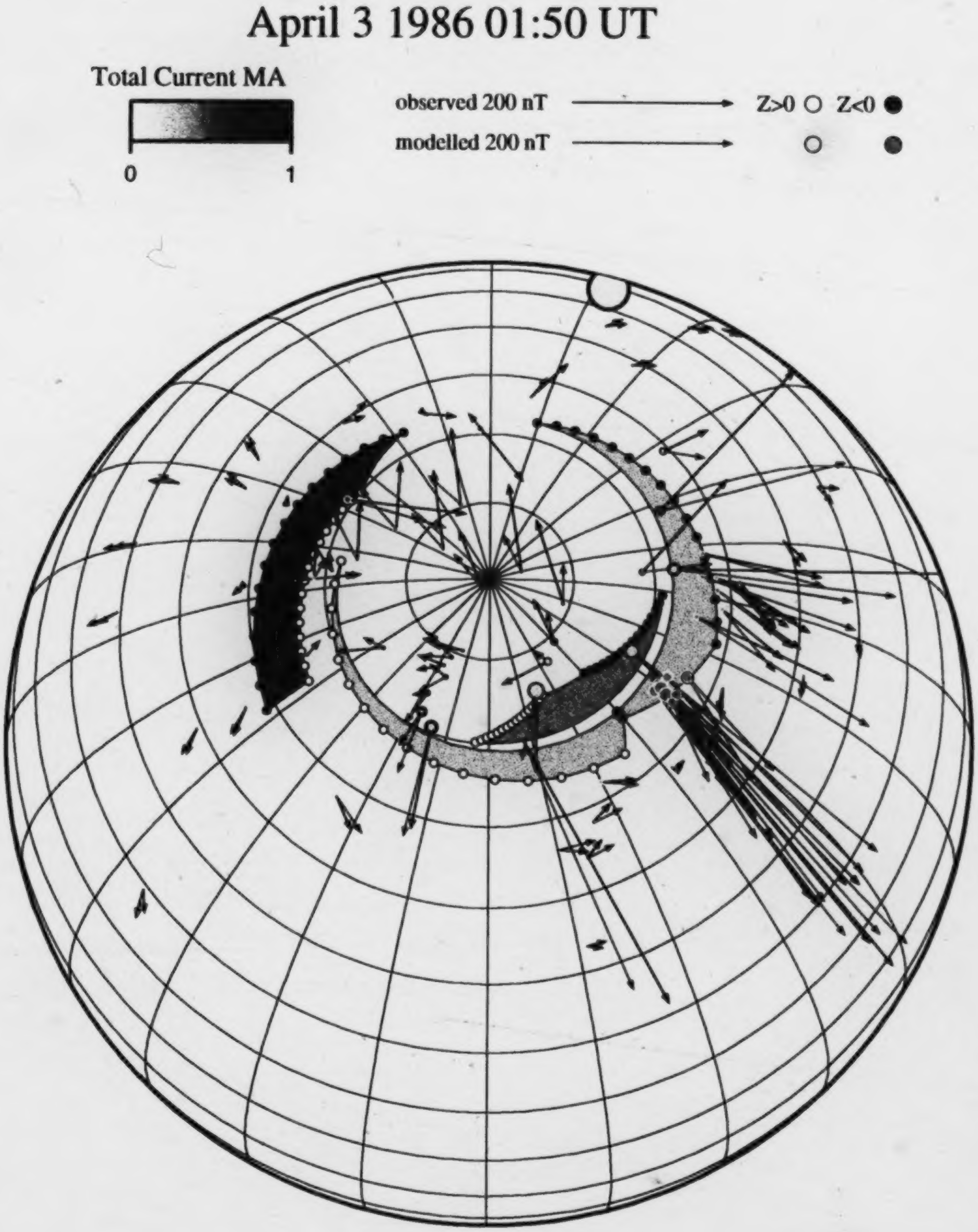

April 3 1986 01:50 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.57 Modelling results at 0155 UT April 3 1986.

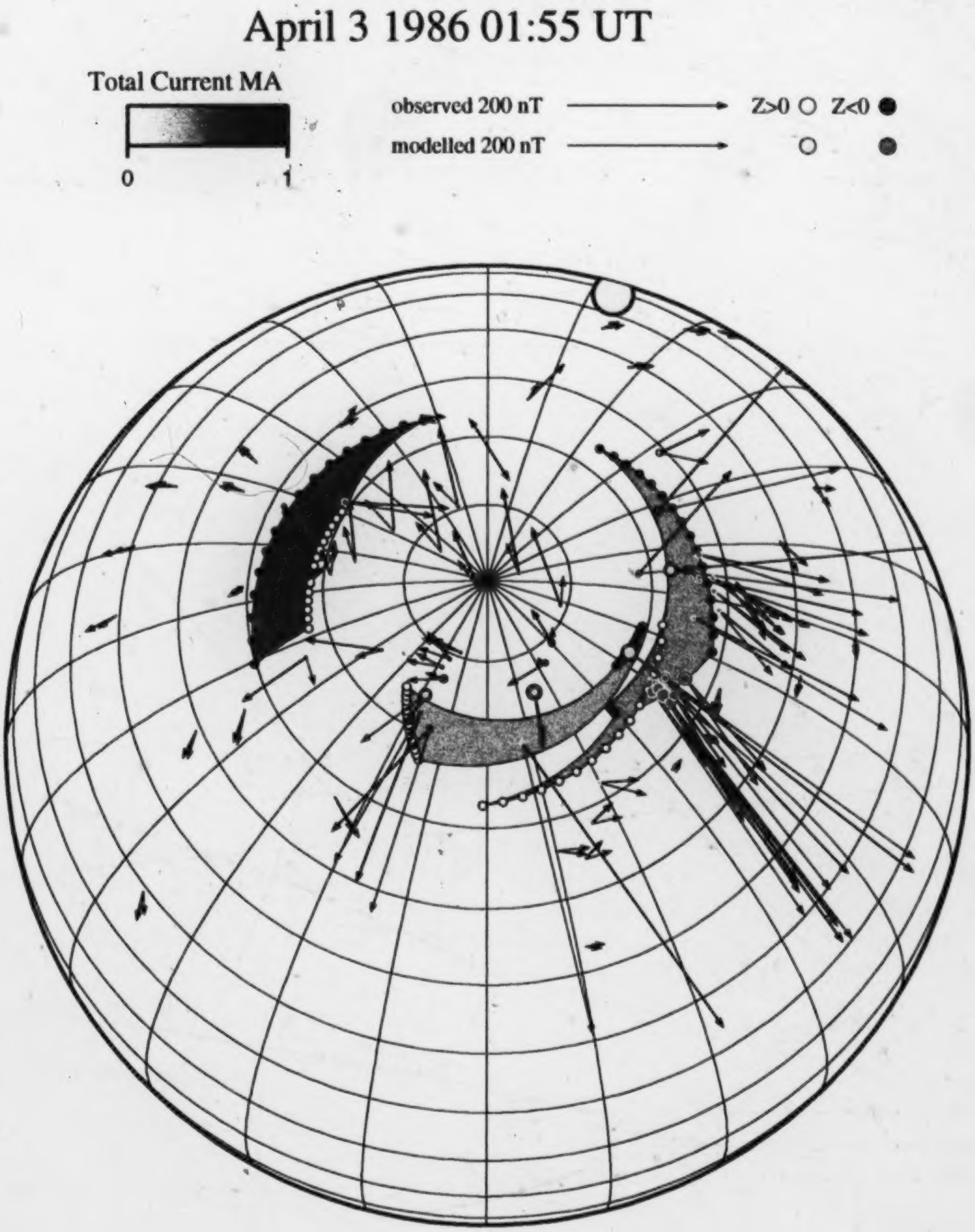
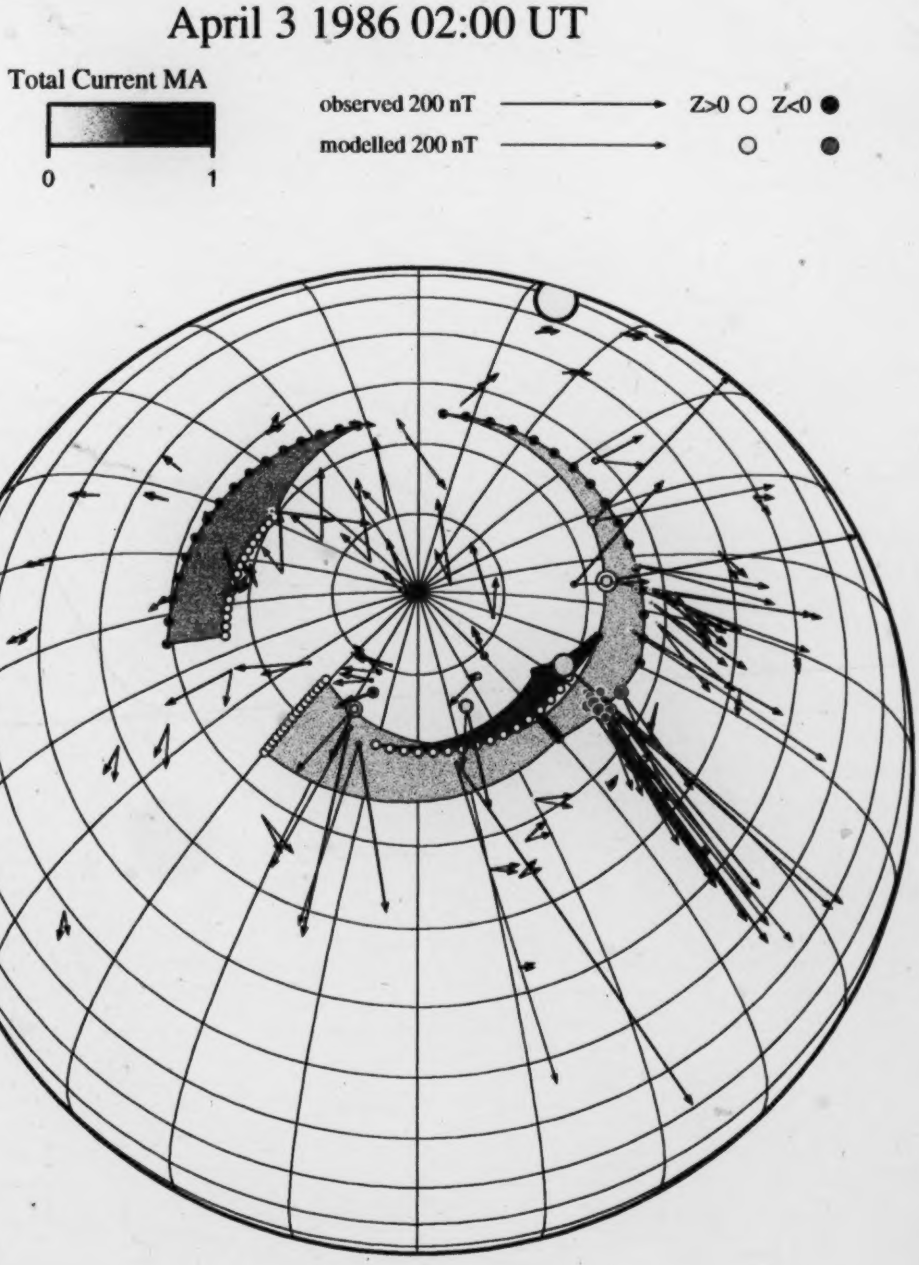

April 3 1986 01:55 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.58 Modelling results at 0200 UT April 3 1986.


April 3 1986 02:00 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●

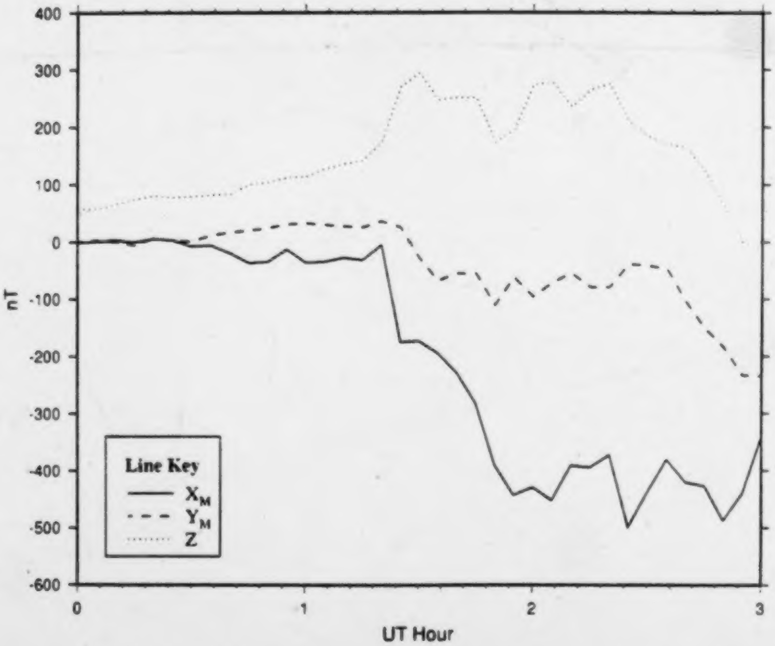
The main difference between Figures 6.53 and 6.54 is an enhancement of the amount of current flowing in the evening sector eastward electrojet and a change in its general layout. The latter has changed so as not to have a generally west to east current flow as in the previous models, with a clear region of downward current in the early evening sector and a clear region of upward current in the late evening sector. In figure 6.54, the current flow within the overall eastward electrojet is generally from west to east, but in the most easterly parts it is locally from south to north. This corresponds to the accepted direction of Region 1/2 current flow in the evening sector. The reason that this general configuration of current has been able to develop is subtle: careful inspection will show that the polar cap stations, especially those nearest the evening sector electrojet, are better represented with the current this way. The model is attempting to add something which was not in it at this stage: Region 1/2 currents which can be responsible for the organized sunward direction of horizontal perturbations in the polar cap. The subauroral perturbations also find a closer match with the current having a large south to north component, and these are relatively heavily weighted. Thus, rather than being a nonphysical result, the reconfiguration of the evening sector electrojet is imparting a strong physical message.

Figure 6.55 shows the same basic configuration as does Figure 6.54, with some reconfiguration of both evening and morning sector electrojets. The latter extends past noon and matches well the noon sector auroral zone perturbation by doing so. Figure 6.56 presents an extreme situation of south to north current flow in the evening sector: this configuration has made a noticeable improvement in the match in the polar cap. The westward electrojet extends north of the eastward electrojet in the evening sector: perhaps since the perturbation at PDB is southward, it should be closer to the evening sector electrojet than shown. Perturbations at SCO in eastern Greenland have resulted in the 'wedge' current approaching that station. As in the previous few model steps, the westward electrojet is fed gradually by distributed downward field-aligned current. Physically, this is more likely to be at the poleward than the equatorward border as shown. A generally good match to the data is obtained at positions along the electrojet. Figure 5.57 is reminiscent of Figure 6.28 in that the westward part of the westward electrojet has been essentially removed from the model by being made small and placed where there are no stations. The 'wedge' is now playing the role of the westward electrojet over Greenland. There is ambiguity in the use of the extra parameters available through having a jointed westward electrojet as is illustrated by the fact that both parts are back in use in Figure 6.58 to model perturbations which are generally similar to those in Figure 6.57. A strengthened southward perturbation at PDB could lead to the argument that the westward electrojet does extend over that station at this time. However, the configuration shown does reproduce the perturbations near PDB well, including those of heavily weighted midlatitude stations.

In this part of the modelling run, more than in that detailed previously, ambiguities have arisen and care must be taken in interpretation. The model appears to be telling us that north-south currents should be included if the polar cap is to be modelled: also a strong current appears to be present at the northern border of the auroral oval.

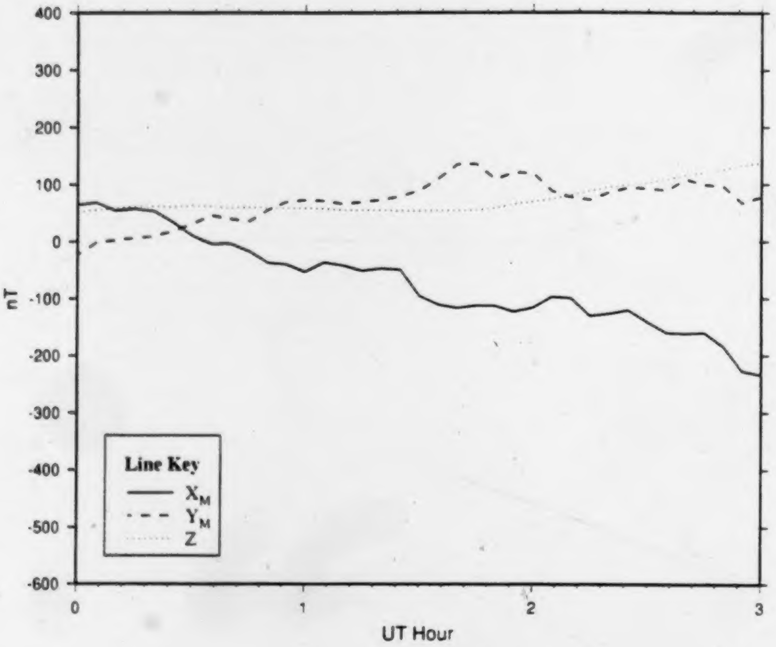
This enhanced current, still referred to here as a 'wedge', is largely called for by perturbations at one station, Bjørnøya, which are shown in Figure 6.59. The steep drop of nearly 200 nT in the X_M component seen at Bjørnøya at 0125 UT is typical of an expansive phase onset. The figure shows 5 minute averages and baselined data supplied by B. Emery [personal communication, 1995]. Observatory records (not shown) with better time resolution were obtained to verify the onset-like nature of the perturbations; such perturbations occur at 0125 and 0150. The rise in Z accompanying the 0125 onset indicates that the ionospheric current was largely south of the station, much as the global modelling indicated. Midlatitude positive Y_M perturbations south of the upward current and negative Y_M perturbations south of the downward current may also be seen through inspection of the superposed data arrows shown on the solution plots (Figures 6.52 to 6.58). In many ways this enhanced northern border current shares characteristics of the substorm current wedge. Further aspects of the onset-like nature of the behaviour of this system will be considered below.

6.59 Perturbations at Bjørnøya on April 3 1986. These are five-minute averaged data.



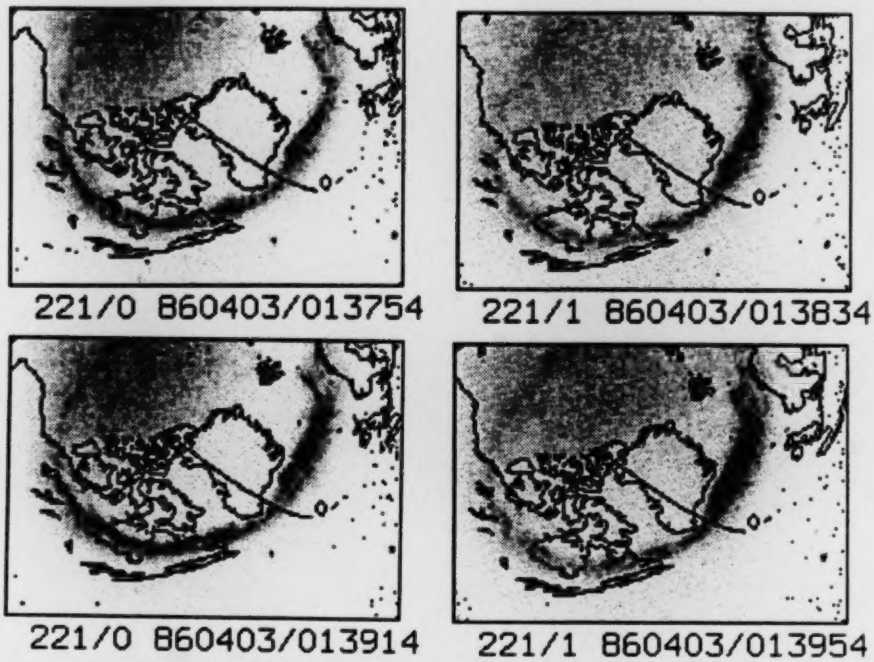
It was noted above, in discussion of modelling in the Russian 155° meridian, that the station at Heiss Island at its northern end initially showed positive perturbations, and later negative. The data from this station are shown in Figure 6.60. The northward component perturbations strengthened at the same time as 'onset' at BJN (0125). Perturbations at this station were generally underrepresented in the global modelling and also likely indicate an enhanced northern border current.

6.60 Perturbations at Heiss Island on April 3 1986.



Further evidence of the high-latitude wedge-like activity is provided in the Viking images of Figures 6.61 through 6.63. Brightness associated with upward currents is seen in the same regions indicated by the modelling results, that is to say in the vicinity of Iceland, and in a form which corresponds to the 'eye' of a substorm. The auroral oval in these images retains its width of about five degrees in the evening sector. In the morning sector, double auroral oval forms. Unfortunately, in agreement with what was deduced above, Bjørnøya is the only magnetic station to which the enhanced northern border of the electrojet comes near. The next closest is Isafjordur in northern Iceland, where a sharp drop in the Z component (about 100 nT) takes place at the onset time of 0130 inferred for this northern border onset. This drop would correspond to a current north of the station: that it is not noticeable at slightly more southerly Icelandic stations confirms that the current involved is quite localized, seemingly in north-south width. The luminosity associated northern border current remains tantalizingly just off the eastern coast of Greenland throughout the time period shown in the images. Thus, the interesting event which has been inferred remains undetected at most magnetic stations. Other lines of evidence for its occurrence must be sought.

6.61 Viking images of northerly onset in the period 0138 to 0140 UT April 3 1986.



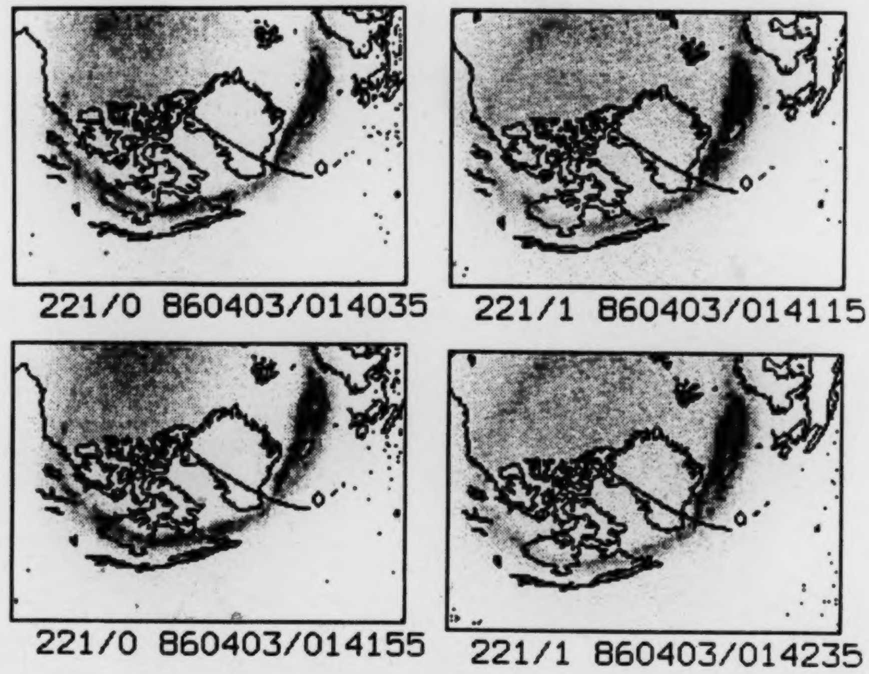
221/0 860403/013754

221/1 860403/013834

221/0 860403/013914

221/1 860403/013954

6.62 Viking images of northerly onset in the period 0141 to 0143 UT April 3 1986.



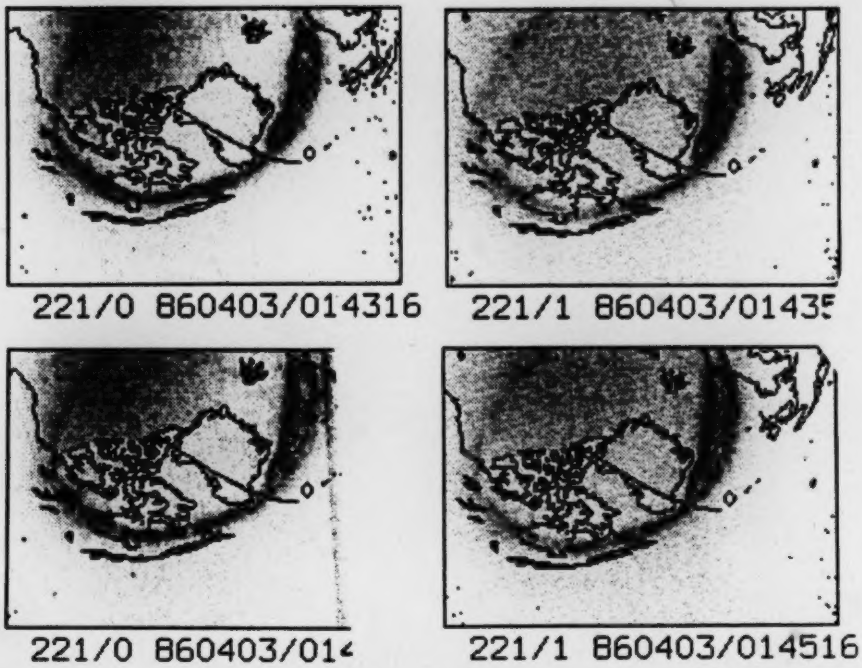
221/0 860403/014035

221/1 860403/014115

221/0 860403/014155

221/1 860403/014235

6.63 Viking images of northerly onset in the period 0143 to 0145 UT April 3 1986



221/0 860403/014316

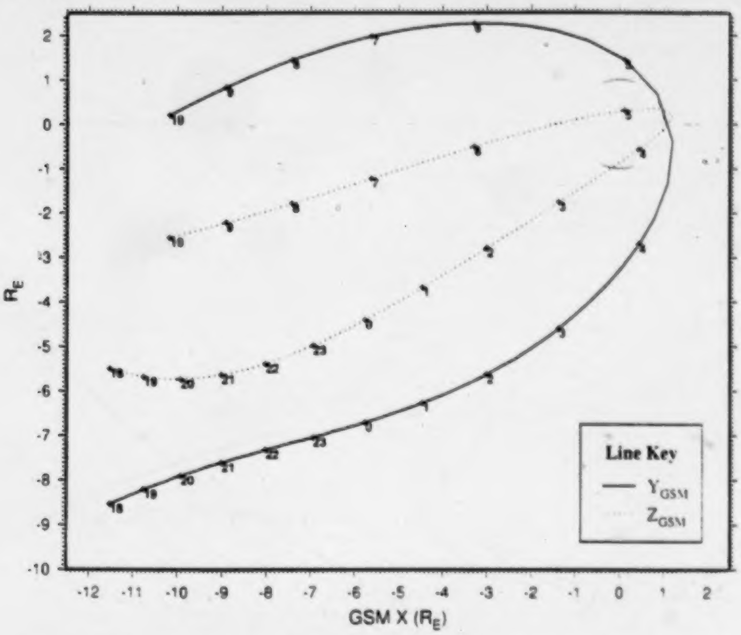
221/1 860403/01435

221/0 860403/014

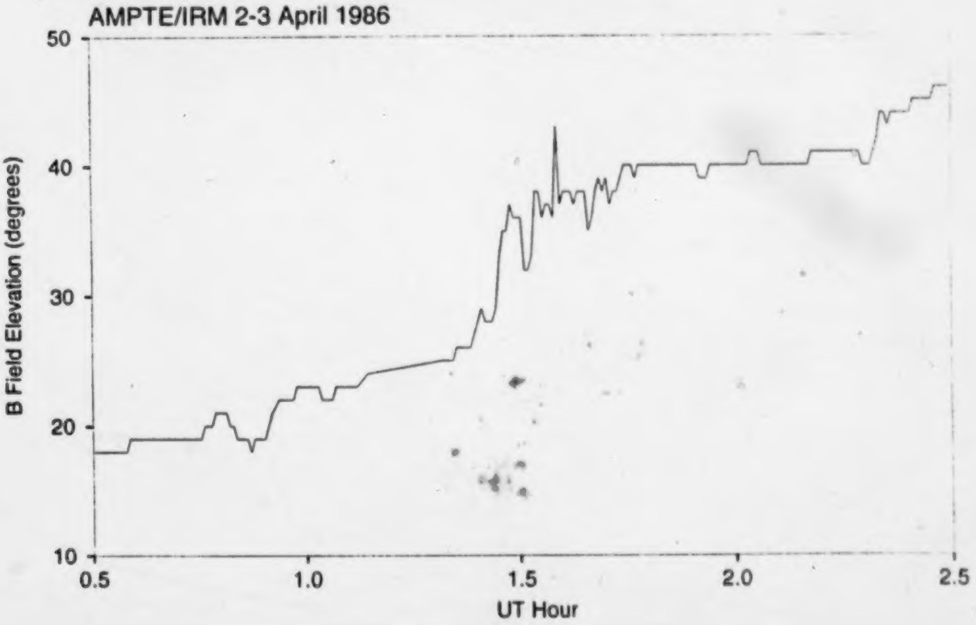
221/1 860403/014516

More evidence that an expansive phase onset took place is supplied by the AMPTE/IRM spacecraft, well situated in the morning sector as shown by Figure 6.64. The spacecraft was approaching Earth from the tail and was south of the equatorial plane during the period of interest. Figure 6.65 shows, through the increase in elevation of the magnetic field, that a clear dipolarization was observed at 0130, superposed on the trend to dipolarization as the spacecraft slowly approached Earth from the tail on its orbit. At the time of dipolarization, the spacecraft was at a radial distance of roughly 7.8 R_E, and at 3 MLT. At the same time, the azimuth of the field also changed from a more downtail to a less downtail value (see Figure 6.66), a signature entirely consistent with dipolarization. The total field was relatively unchanged, however (Figure 6.67), which indicates that the spacecraft likely did not cross some magnetospheric boundary. The abrupt increase in proton bulk speed shown in Figure 6.68 may correspond to a proton injection as seen at AMPTE/IRM, or could correspond to beams in the plasma sheet boundary layer. This speed increase and the associated proton temperature increase shown in Figure 6.69 (the upper temperature shown, 2×10^8 K, corresponds to about 18 keV) are perhaps best interpreted as indicating that a dramatic change occurred at 0128 as opposed to specifying what it was. The proton density (not shown) did not dramatically change around 0130 and a gradual increase observed after that time is likely best explained through the approach of the spacecraft toward Earth.

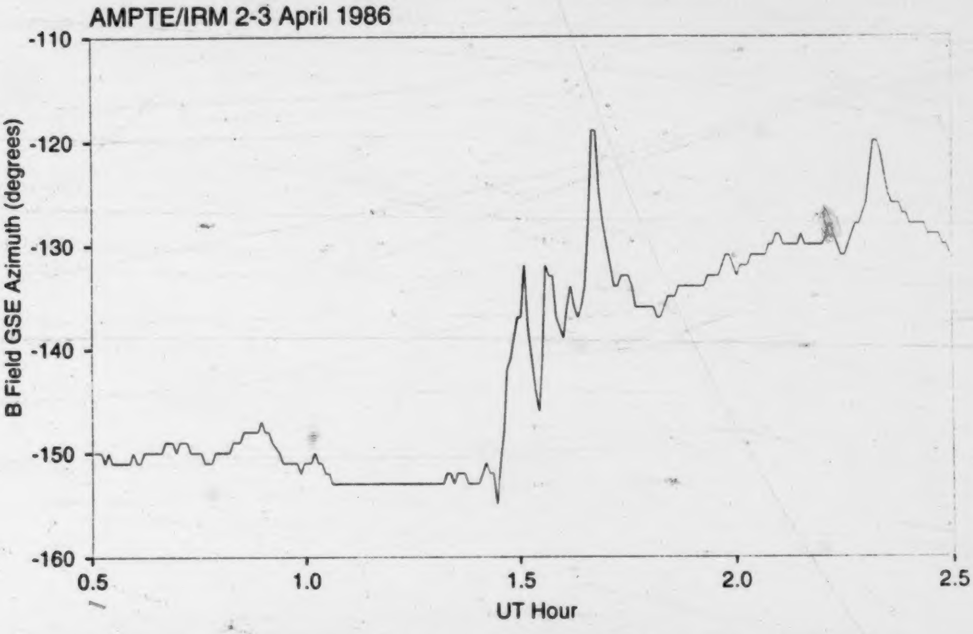
6.64 Position of AMPTE/IRM satellite on April 2-3 1986. GSM (X,Y) pairs form solid line, (X,Z) pairs form dotted line. UT is indicated beside crosses each hour.



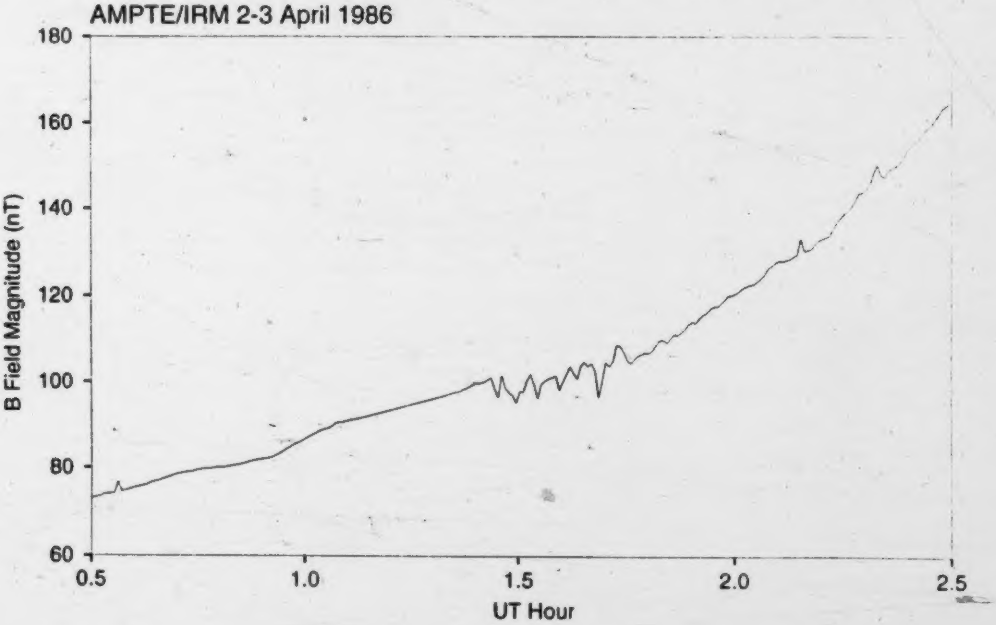
6.65 Elevation of the magnetic field at AMPTE/IRM on April 3 1986. Prior to 0130 (1.5 UT) elevation is tail-like; after this time more dipolar.



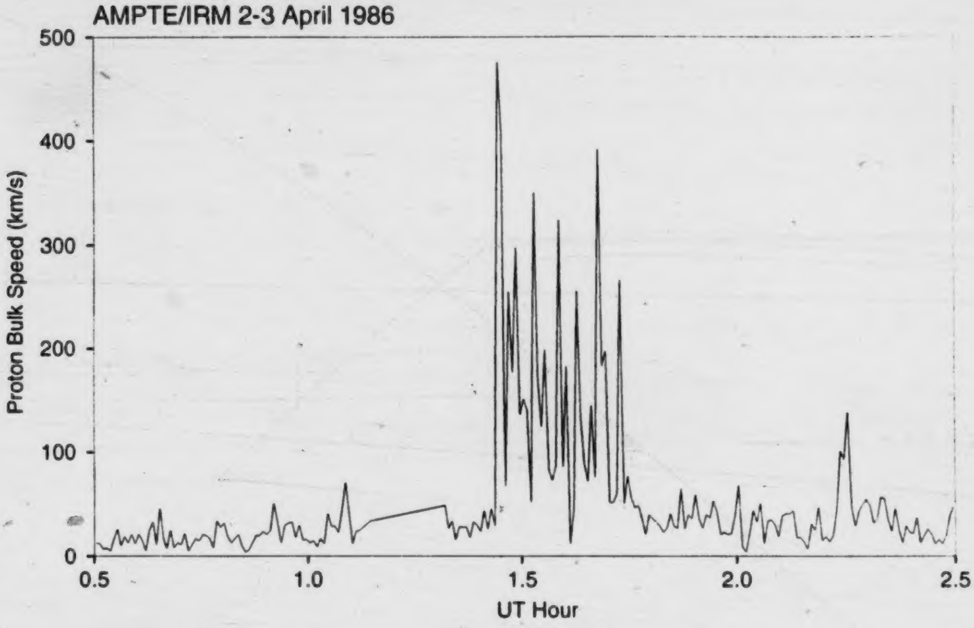
6.66 GSE azimuth of the magnetic field at AMPTE/IRM on April 3 1986. Change away from -180° antisolar angle corresponds to dipolarization.



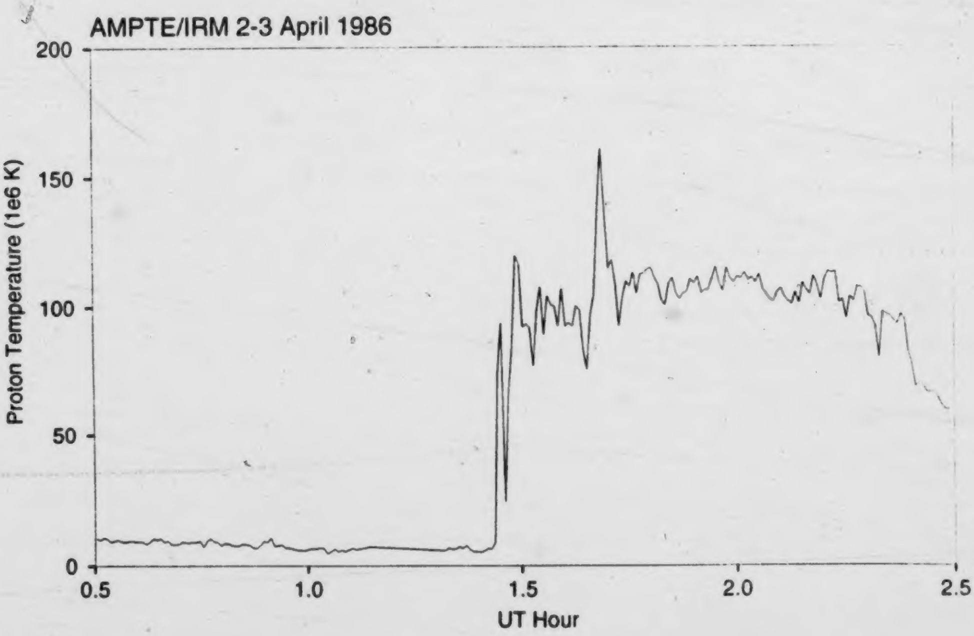
6.67 Magnitude of the magnetic field at AMPTE/IRM on April 3 1986. Changes around 0130 UT are small, superposed on the increasing trend approaching perigee.



6.68 Proton bulk flow speed at AMPTE/IRM on April 3 1986.



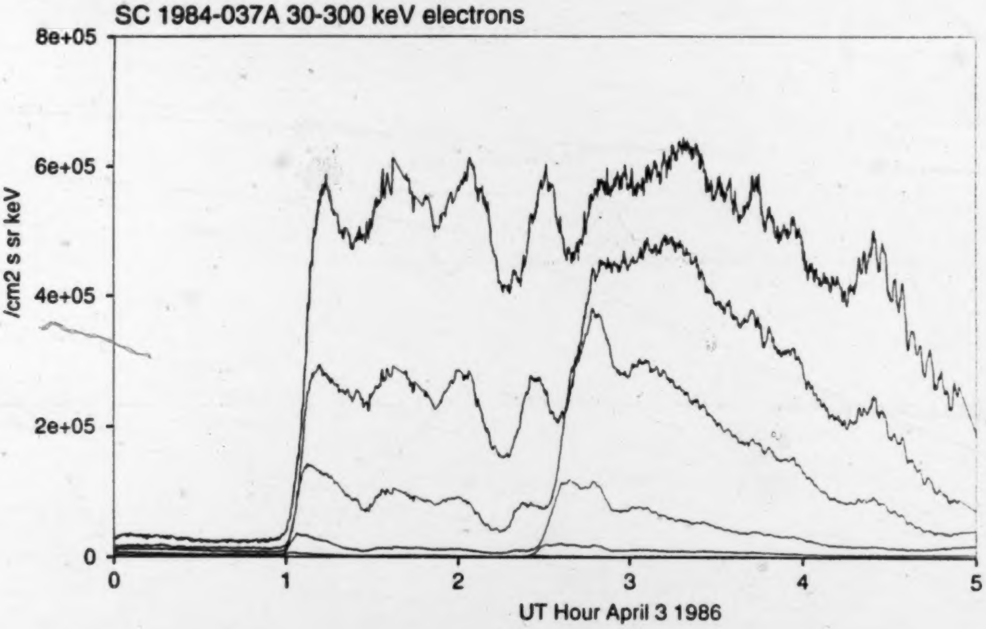
6.69 Proton temperature at AMPTE/IRM on April 3 1986.



While the data presented here (ground signature, dipolarization, injection at synchronous orbit (see below), possible injection as seen at AMPTE/IRM) support the claim that an actual expansive phase event occurred at high latitude in the morning sector, Opgenoorth *et al.* [1994] have discussed this activity and classify it as 'intense morning-sector precipitation', which they take to be a newly-discovered feature of the substorm recovery phase. They claim that the recovery phase is that of the substorm which began at 0100 (here considered to have commenced at 0055). While they consider effects of the activity at Leirvogur and in Greenland, they claim that no effects are seen at Kiruna. To consistently explain this, one should note that activity of this sort, producing a negative X_M bay at Bjørnøya, will produce only minor perturbations as much further south as Kiruna, at least in the X_M component. The reasonable fits in the modelling shown here, in a well-constrained area like northern Scandinavia, indicate that the surface magnetic field is well modelled by the wedge plus morning sector electrojet, with the latter dominating over the Scandinavian mainland. The Z component is well modelled (see diagrams early in the sequence in which Z changes sign at Soroya) although it is consistently slightly too large in magnitude where negative. Although Opgenoorth *et al.* [1994] claim that no particle injections are seen associated with this event, the situation is not entirely clear. Since the purported onset occurs at high latitude, injections might not necessarily be seen at synchronous orbit [Kamide and McIlwain, 1974]. However, the SC 1984-037A spacecraft does have multiple 'bumps' in its electron detector at times which could be associated with the event as shown in Figure 6.70. Expansion of scale from comparable Figure 6.21 shows electron injections at 0055, 0125, and 0150 UT April 3, corresponding to expansion phase onsets at high latitudes. Dispersed onsets at 0215 and 0230 are also seen. SC 1982-019A has, in noisy data (refer back to Figure 6.22) a proton peak at the right time as might be expected due to its location west of the likely injection point. This spacecraft also has a broad but little-dispersed electron peak at about 0120 UT: this is too early, and the spacecraft is in the wrong location, for the event under discussion. As discussed in the previous section, this signature could be that of a second injection associated with the 0055 onset, possibly due to the flare-up and 'wedgelet' in the vicinity of Poste-de-la-Baleine. The claim that a particle injection is detected at AMPTE IRM is reinforced by the fact that the spacecraft is at an L value of 9.15, considerably greater than that of a synchronous spacecraft (which is always near L=6.6).

At about 0135 UT, the ISEE spacecraft, also located in the morning sector at Y=-12, but well downtail at X=-18.3, detected a large increase in flux of 6 keV electrons and protons (by a factor of 1000 for electrons and about 5000 for protons - data not shown). This is roughly the energy band in which AMPTE/IRM also detected a large temperature increase. The ISEE data was already shown in an illustrative fashion in Chapter 1. Since (see Figure 1.17) the overall particle density did not increase by a factor exceeding three, the most likely explanation is a traversal of the spacecraft by the lobe/plasma sheet boundary, which separates cool, low-density lobe plasma from the higher energy and denser plasma sheet particles. At this time, the magnetic field started to display high frequency irregular variations (visible in Figure 1.16) and the Z component increased. Relating these changes to the 0055 onset is problematic: they are more likely related to the 0125 event and support its nature as an 'onset'. A time delay from onset to ISEE of under 10 minutes corresponds well to the propagation velocity found by Jacquey *et al.* [1993] for tailward propagation of cross-tail disruption signatures associated with substorms. In this regard it might be noted that the Z component shows about a 5 nT increase (i.e. becomes more dipolar) at the time of the particle signatures (data not shown).

Figure 6.70 Electron flux at satellite SC 1984-037A April 3 1986. Linear scale with same energies as those of Figure 6.21.

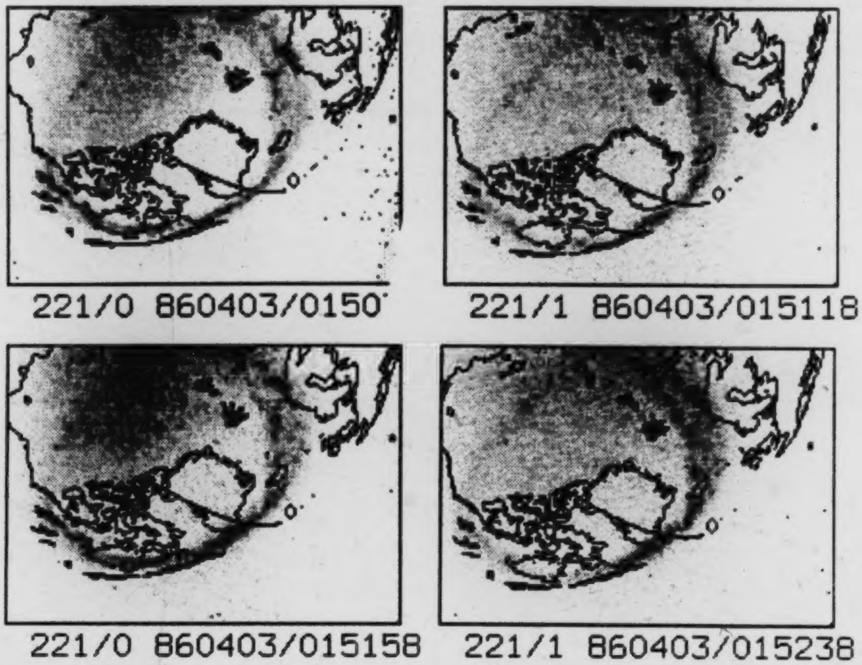


It is important to differentiate the behaviour at the northern edge of the electrojet, as described above, from poleward motion which can occur at expansive phase onset (but which does not appear to accompany the onset at 0055). Such northward motion of the northern boundary of the electrojet is described by Kisabeth and Rostoker [1974]. In that study, by tracing the extrema of the Z-component perturbation along a latitude profile, a steady progression northward of about 6° in twelve minutes was seen. Time resolution was 2 s in the original data and a plot of contours in the Z component-time plane with high time resolution showed smooth poleward progress of the northern border, while the southern border remained approximately fixed at roughly 63°. In the present case the poleward expansion occurs well after the major onset of 0055 UT, not immediately following it. Magnetic data are too sparse for use in attempting to follow poleward motion in this case; however, the images do not suggest a large amount of poleward motion. An event similar to that analysed here was studied by Hones *et al.* [1987] using ground magnetograms, Viking images, and particle data at synchronous orbit as near-Earth data, and with tail data at 17 R_e from ISEE 1. The event was during the latter part of the CDAW 9C period (see Chapter 8). The poleward expansion was by about 6° in six minutes, and at the more northerly stations (for example Barrow) the expansion corresponded to an onset-like signature in the H component. As in the present case, the particle data showed multiple flux enhancements, one of which was at the time of initiation of northward motion. In the Hones *et al.* study, the activity spans, eventually, the local times 0300 to 2100, that is from the morning to the evening.

There have been references above to onset-like signatures at satellites at 0150 UT. On the ground at this time, a distinct onset is seen at Greenland station Frederikshåb (FHB) at 72° invariant latitude, while slightly southward Narssarssuaq shows only an intensification in electrojet strength after about 0200 UT (see Figure 6.24). The main electrojet is south of both stations as evidenced by a positive Z component (see Figure 6.71). At the onset time a small positive X bay is seen at Iqaluit accompanied by a negative shift in the Y component (see Figure 6.14). The latter is consistent with an upward current south of Iqaluit. At Poste-de-la-Baleine (see Figure 6.43), a negative H bay occurs but not much is seen in the D, which would indicate that the upward field-aligned current associated with this wedge is east of that station. This onset is also visible in Viking images as a sudden brightening in the region of Frederikshåb, as seen in Figure 6.72 (lower left). At about 0155 UT the next station north in the Greenland west chain, Godthåb, sees a brief onset also. Thus this activity is definitely to the north of the main westward electrojet.

Figure 6.71 Perturbations at West Greenland station Frederikshåb, 72° invariant latitude. These one minute data show an onset at 0055 and another distinct onset at 0150 as evidenced by a decrease in the H (solid) and Z (dotted) components.

Figure 6.72 Viking images showing sudden brightening in West Greenland.

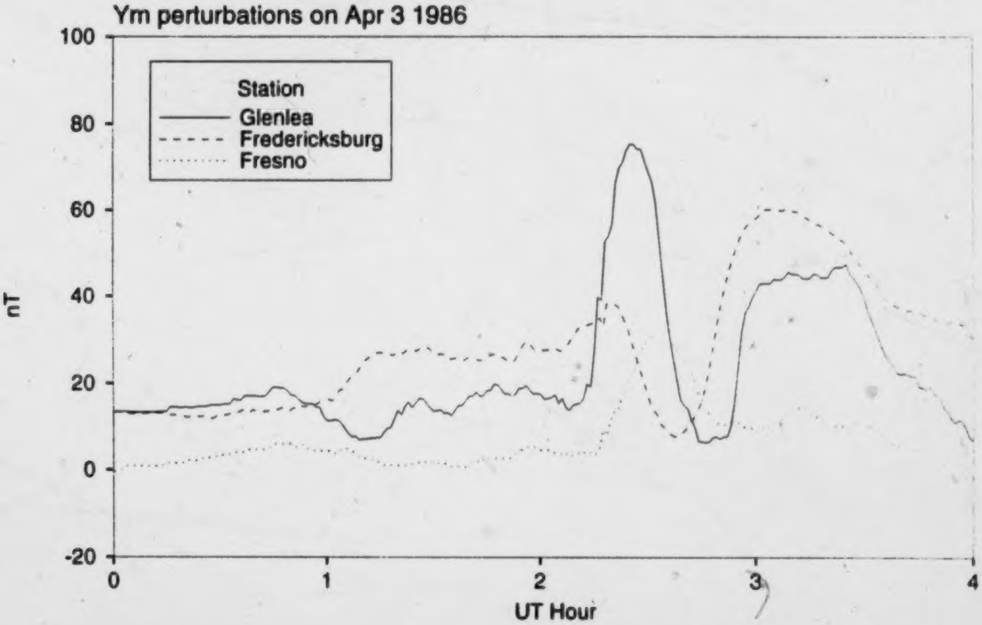


g. Second major onset

During part of the period from 0200 to 0300 UT there is evidence of a very strong substorm current wedge. A second current wedge is active at higher latitude and possibly to the east of the major wedge during the latter part of the interval. Viking images are not available, however much of the magnetic activity occurs over North America. Thus the wedge can be well described and modelled from magnetic data. Dynamics Explorer conjugate data are available, but with much lower spatial and temporal resolution than Viking's, they are of mainly qualitative use. During this period, the IMF field is solidly southward - the onset at roughly 0215, much like that at 0055, shows no strong evidence of triggering by the interplanetary magnetic field. This would appear to speak against mechanisms, such as that described by Lyons [1996], which depend upon northward turnings as a trigger for substorms.

Midlatitude stations in North America give evidence of the major substorm current wedge during this period (see Figure 6.73), and indicate that the activity has moved well into the evening sector. The positive Y_M perturbation which peaks at 0220 at Glenlea indicates upward current flow in roughly the 1800 MLT local time sector. The positive Y_M feature is followed by a return to small or near-zero values with some aspects of a transition bay [Rostoker, 1966]. The Y_M component by 0300 has climbed to a higher value than before the passage of the wedge indicating the 'placement' of a region of upward field-aligned current north of Glenlea. Fredericksburg, further east, sees a marginal transition bay, with an initial increase in Y_M at about 0215 followed by a decrease. Upward field-aligned current was already west of the station at the time of formation of the wedge, but during most of the lifetime of the wedge, downward field-aligned current dominated this longtiude sector, resulting in a primarily negative perturbation in the Y_M component. The rise of Y_M to higher values than before passage of the wedge occurred at about 0250 UT, before the corresponding rise at Glenlea. This suggests that intrusion of the westward electrojet into the evening sector took place, with a predominantly upward field-aligned current regime in the auroral zone of eastern and central North America. At Fresno, near the west coast, primarily the effects of upward field-aligned current were seen while the wedge was active, and the Y_M component returned to near its initial values after the wedge had died off. The region of upward field aligned current emplaced north of eastern and central North America was not seen in the west.

Figure 6.73 Y_M Perturbations at Glenlea, Fredericksburg, and Fresno, April 3 1986.



With the information provided by analysis of the midlatitude perturbations a fair idea is obtained of the placement of the substorm current wedge. Initial conditions roughly corresponding to the wedge at about 0220 were formulated and used to solve for the currents during the hour 0200 to 0300 under consideration. The run displayed here was started from low-χ^2 results of that run. That expected results generally consist of a westward motion of the region of most activity is seen in motion of the brightest region toward the west in DE images of the southern hemisphere (not shown). Figure 6.74 shows the initial conditions used in this run, and Figures 6.75 through 6.86 show modelling results at five minute intervals from 0200 to 0255 on April 3, 1986.

Figure 6.74 Initial conditions for modelling during the hour 0200 to 0300 UT April 3 1986. Stations are shown for reference as grey circles, from east (clockwise): SOR, LRV, PDB, FCC, YKC.

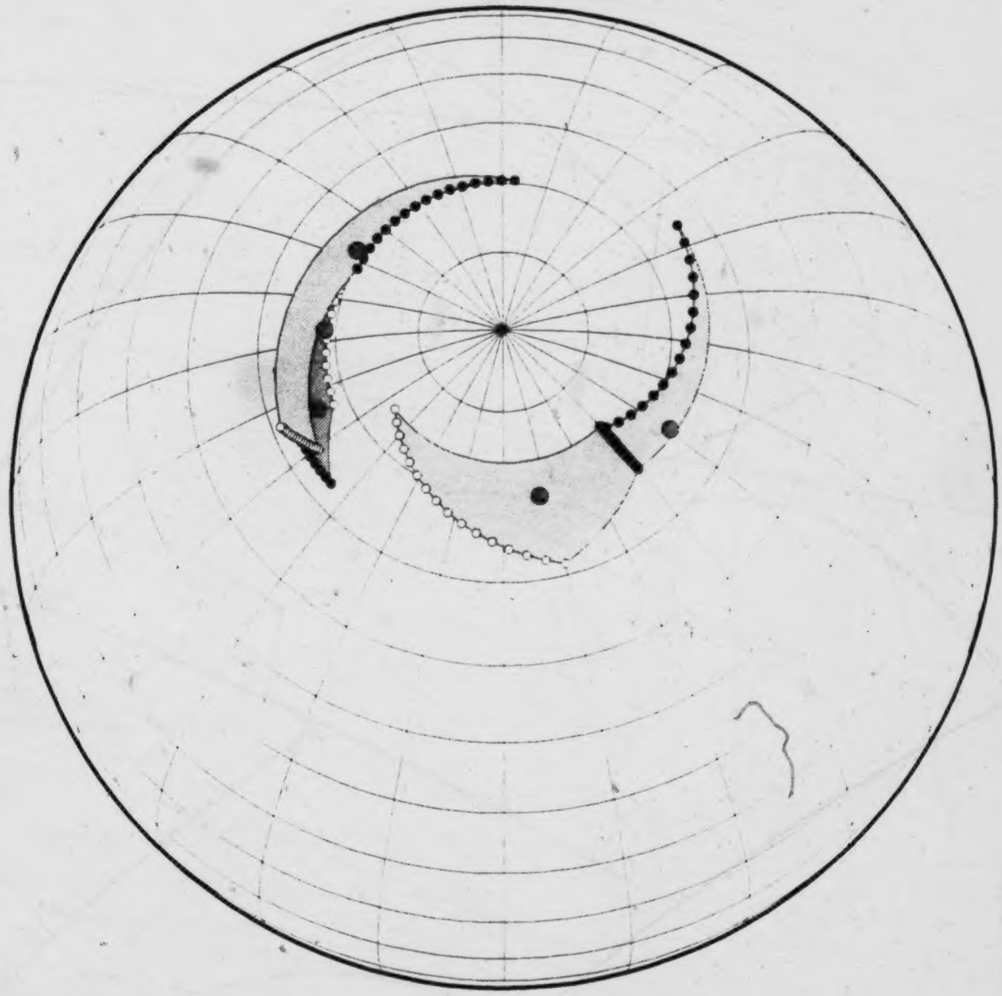


Figure 6.75 Modelling results at 0200 UT April 3 1986.

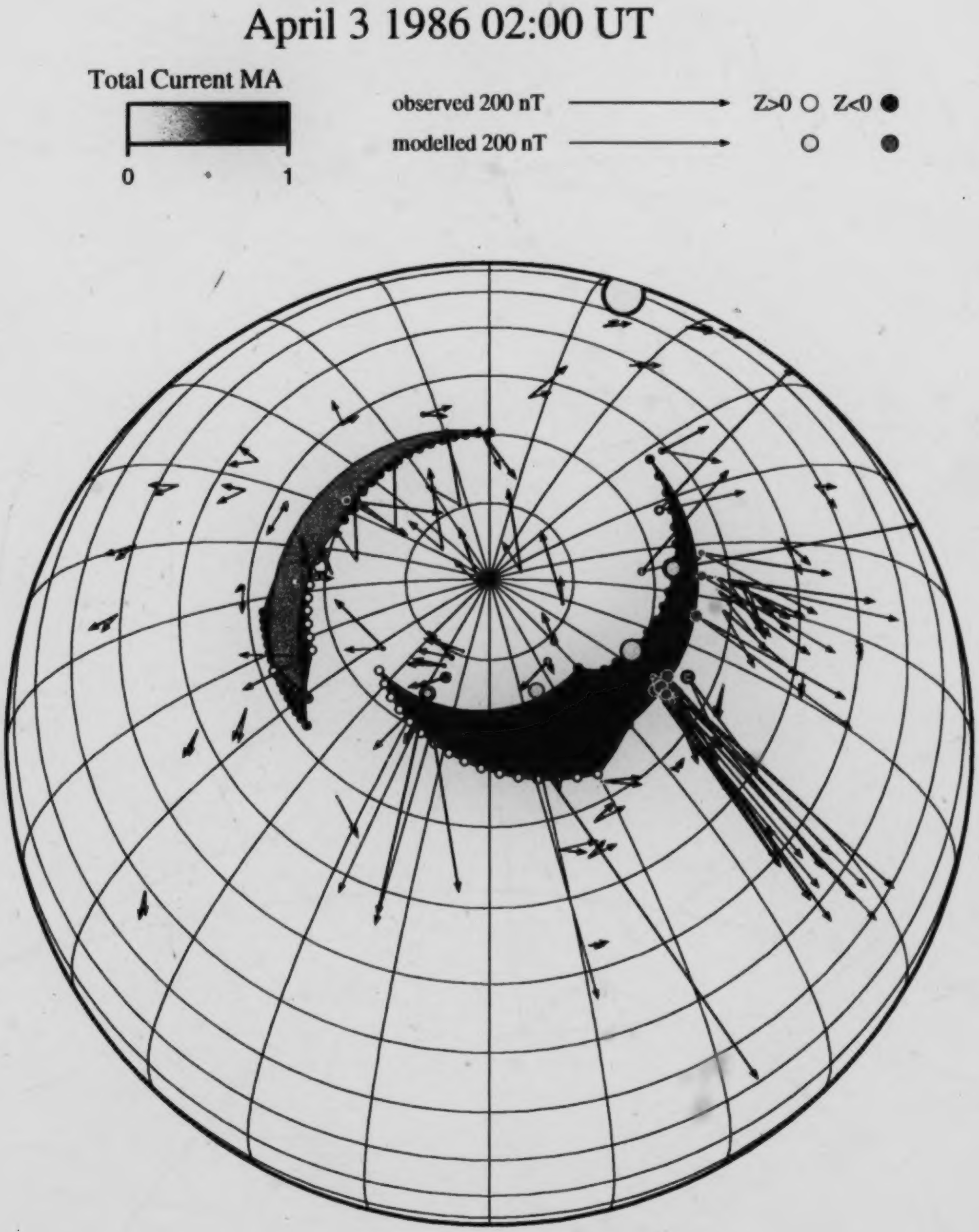

April 3 1986 02:00 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.76 Modelling results at 0205 UT April 3 1986.

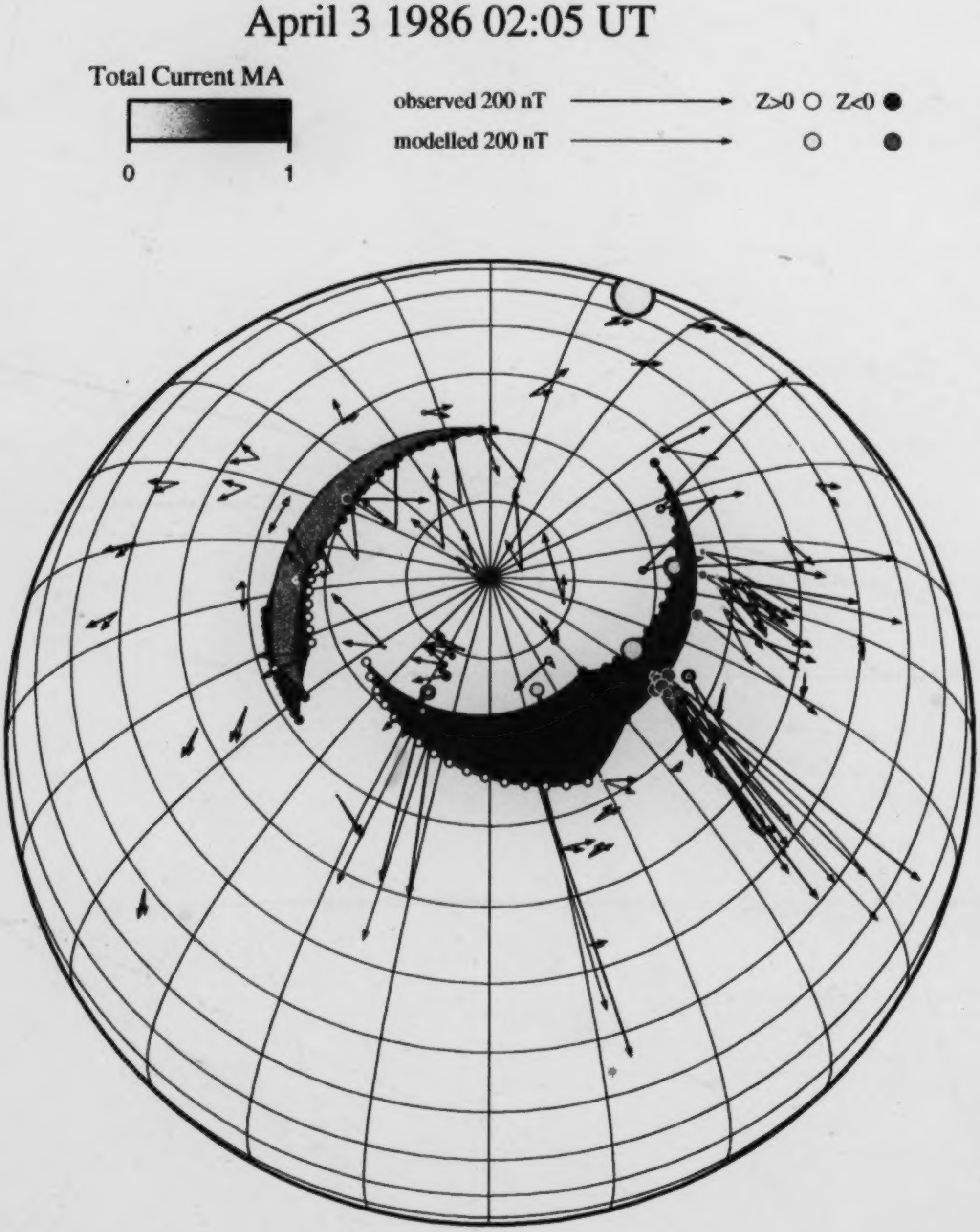

April 3 1986 02:05 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.77 Modelling results at 0210 UT April 3 1986.

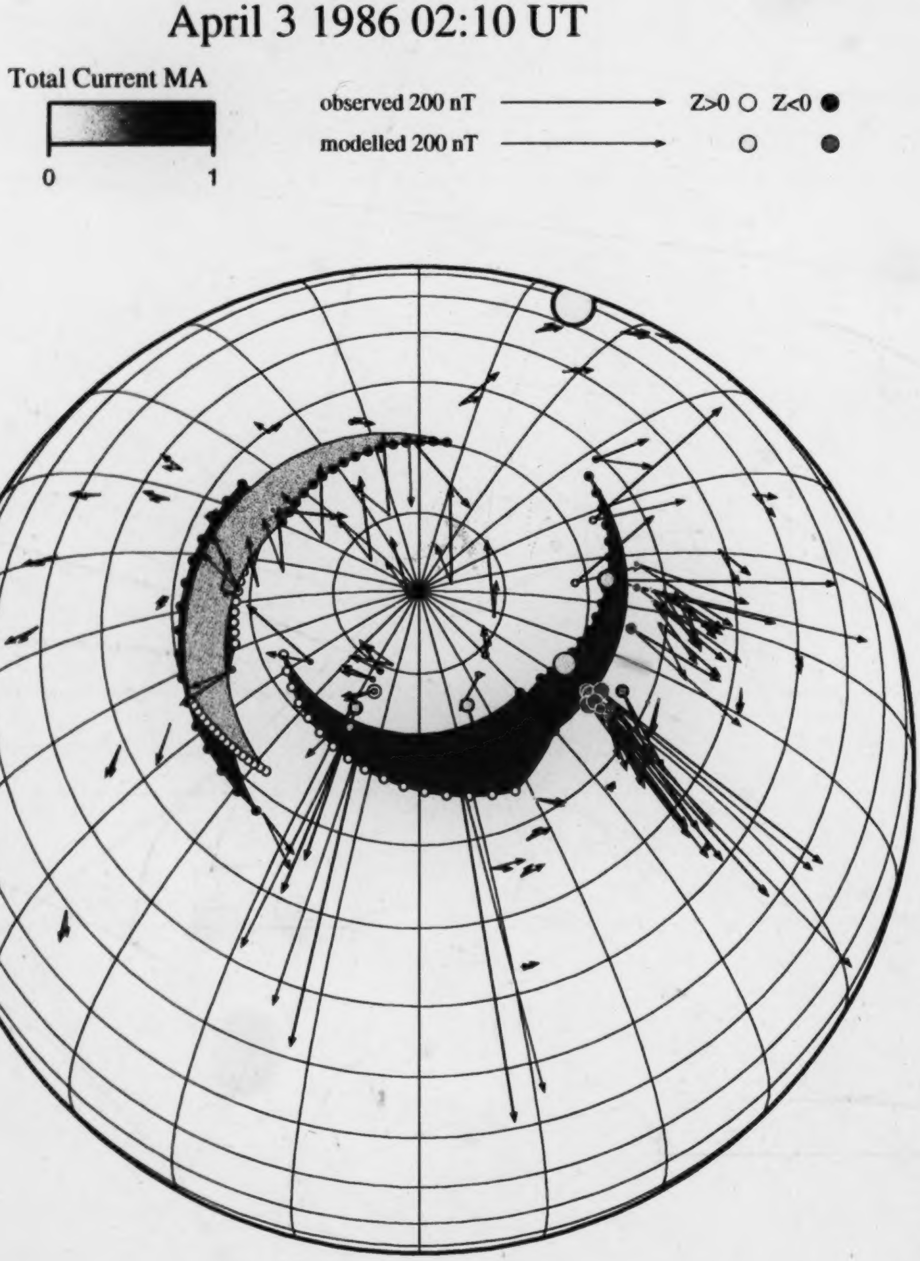

April 3 1986 02:10 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0
Z<0

Figure 6.78 Modelling results at 0215 UT April 3 1986.

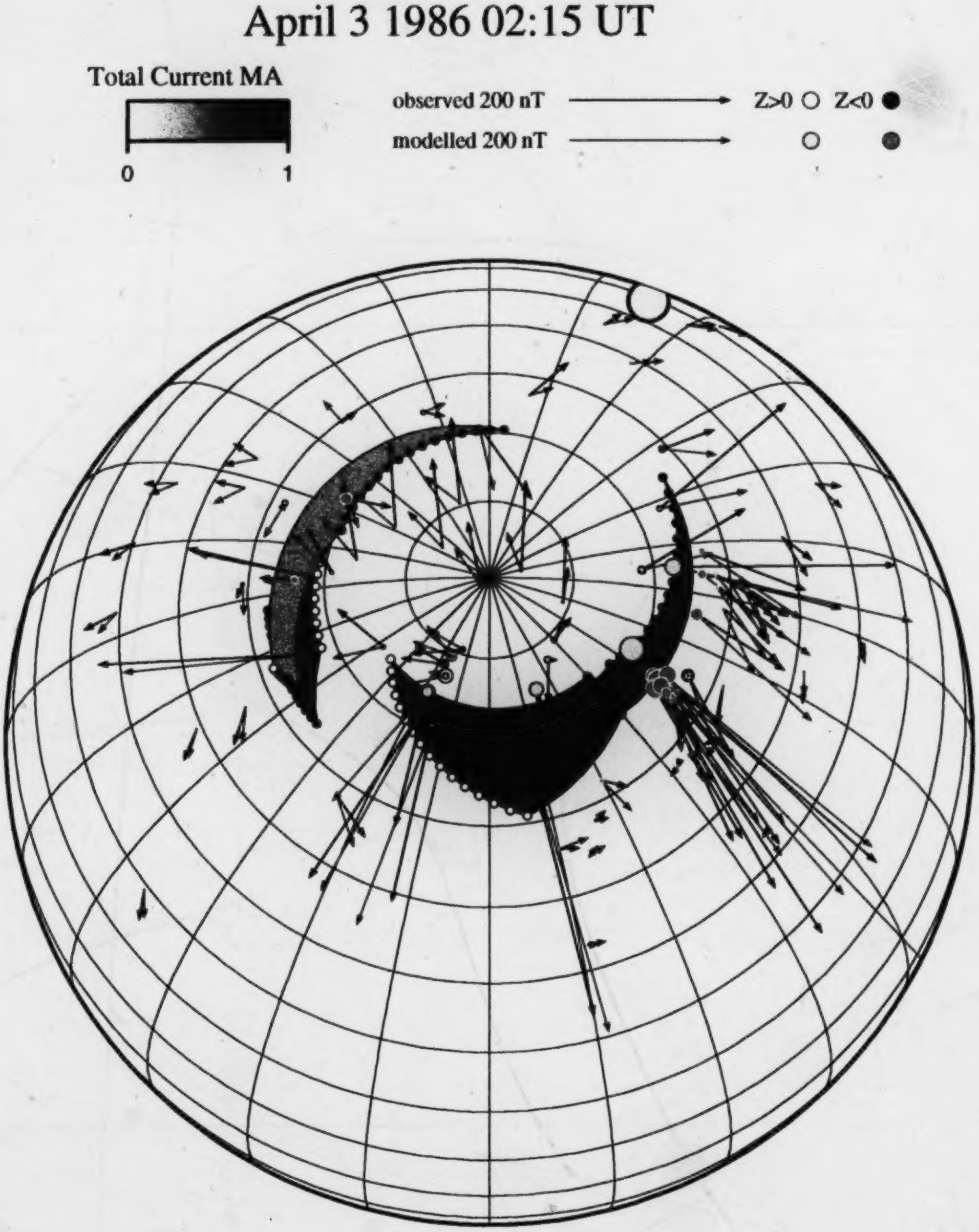

April 3 1986 02:15 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.79 Modelling results at 0220 UT April 3 1986.

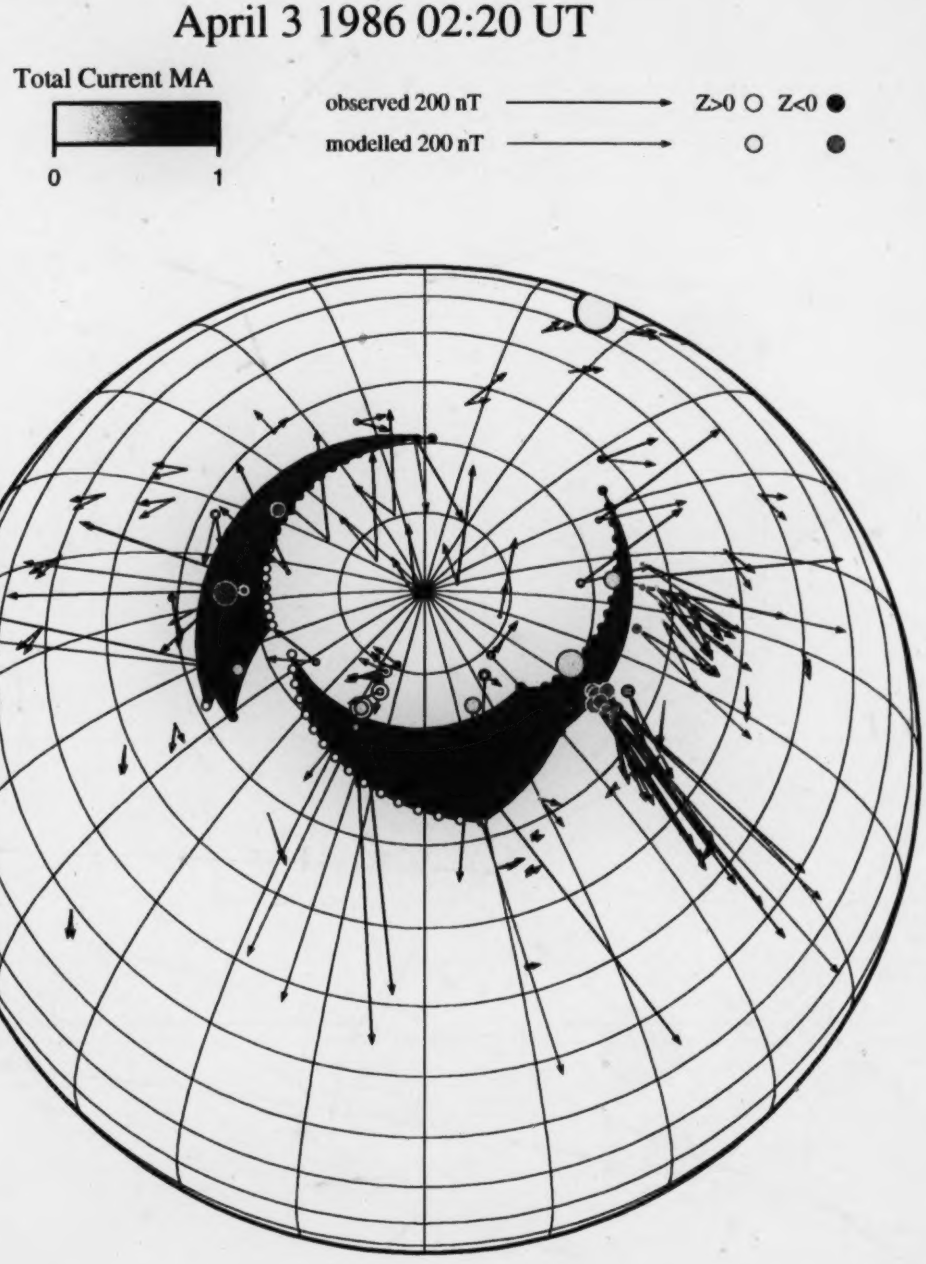

April 3 1986 02:20 UT
Total Current MA
0
1
observed 200 nT
Z>0 ○ Z<0 ●
modelled 200 nT
○ ●

Figure 6.80 Modelling results at 0220 UT April 3 1986 including Region 1/2 currents.

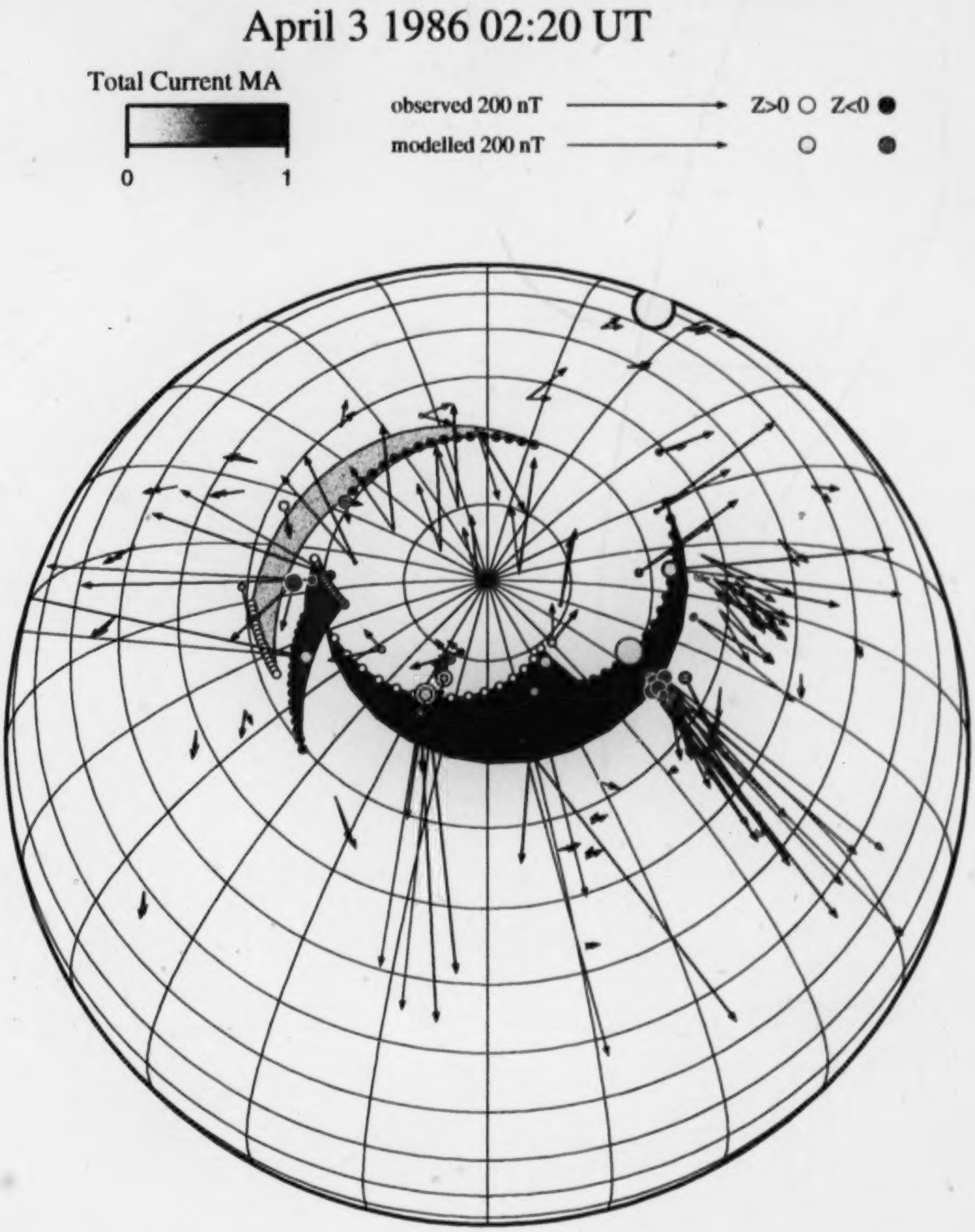

April 3 1986 02:20 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.81 Modelling results at 0230 UT April 3 1986.

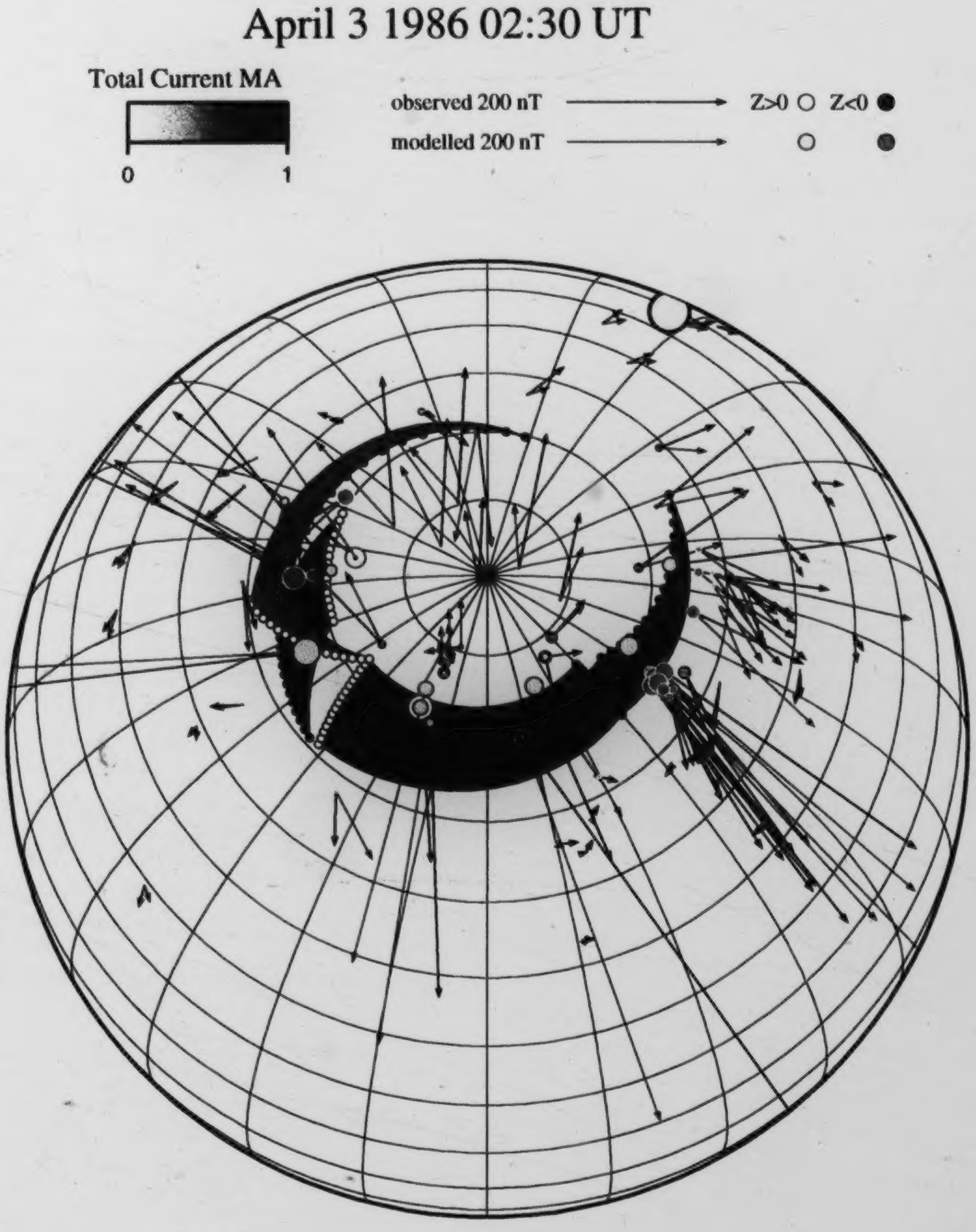

April 3 1986 02:30 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●

Figure 6.82 Modelling results at 0230 UT April 3 1986 including Region 1/2 currents.

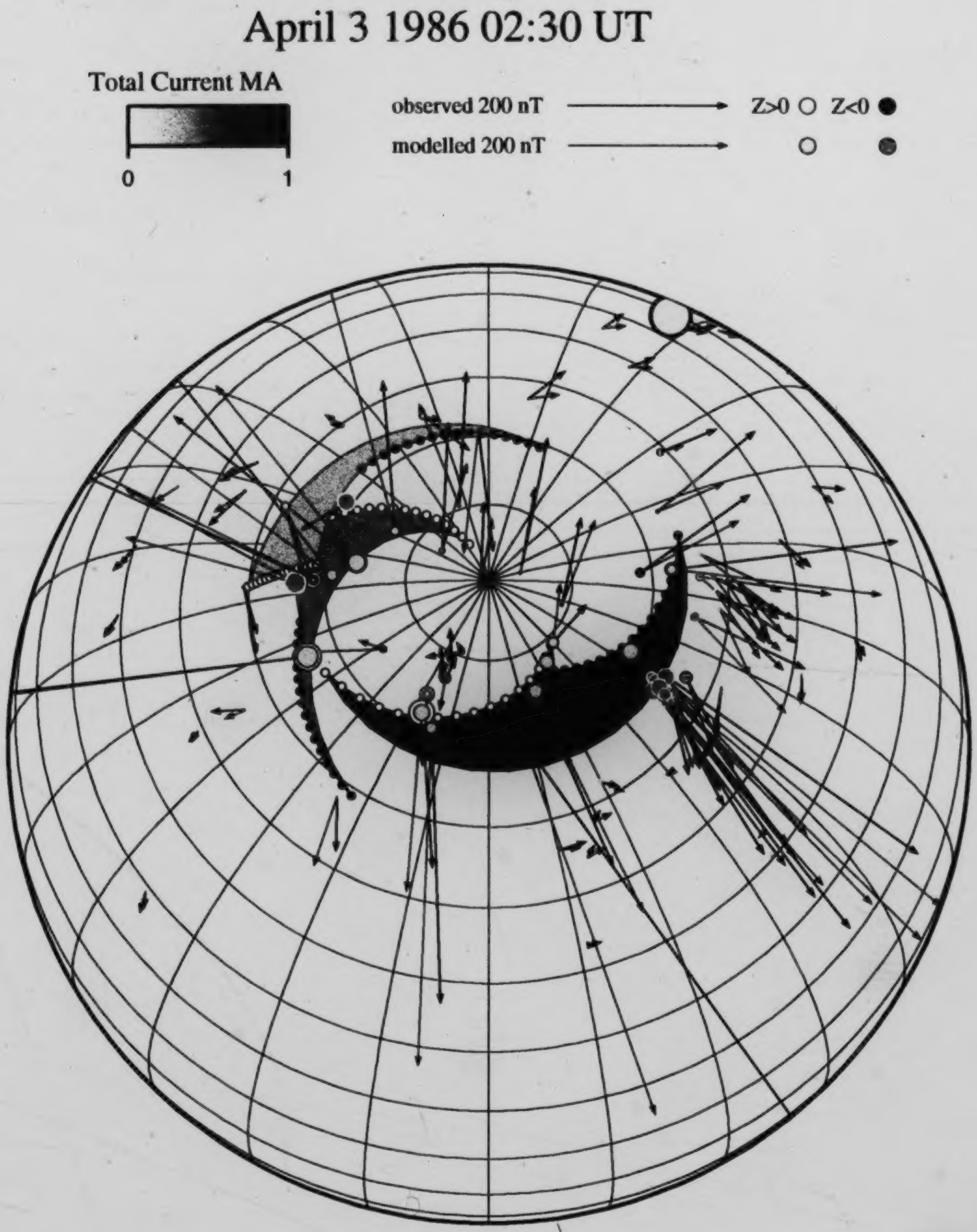


Figure 6.83 Modelling results at 0240 UT April 3 1986.

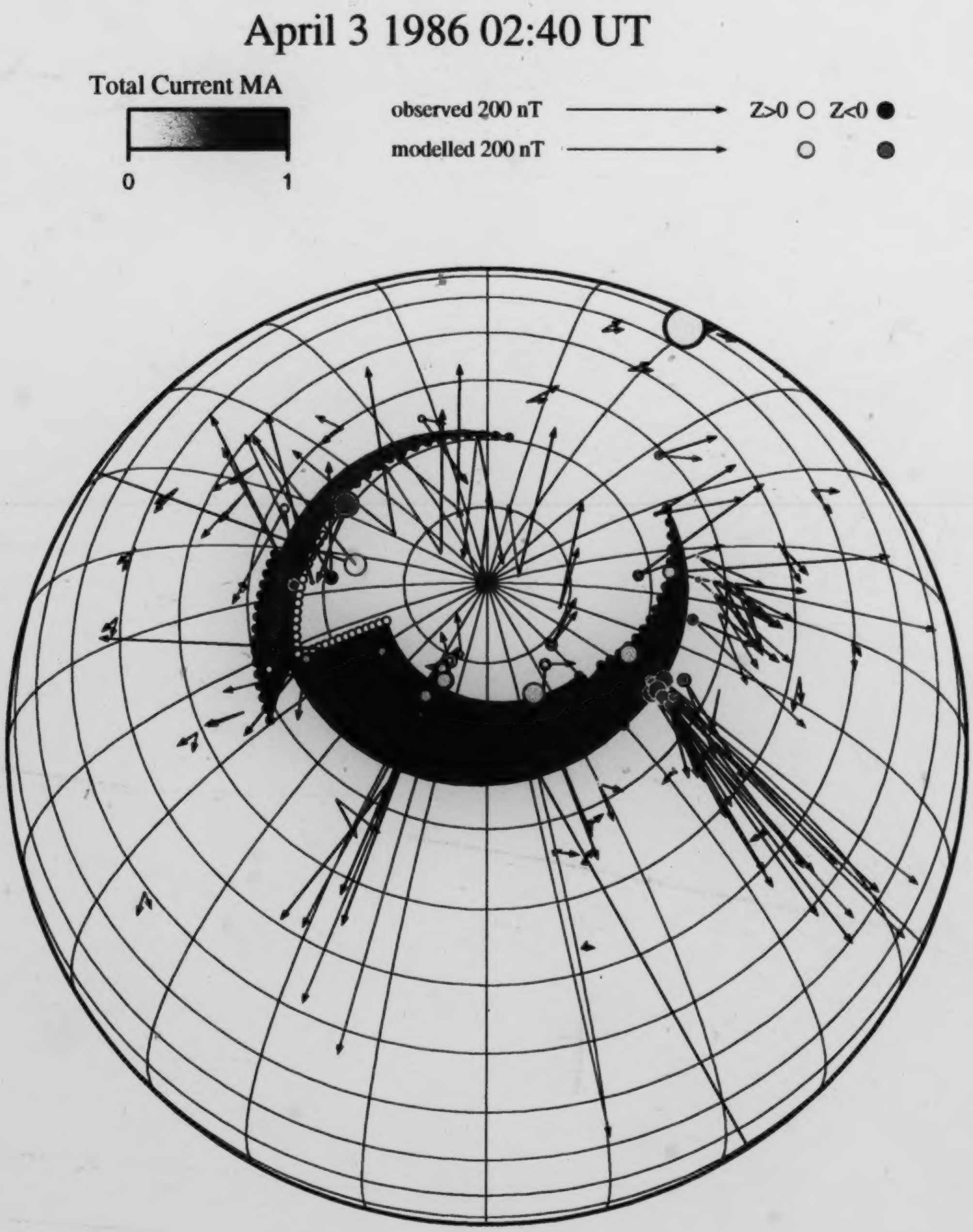

April 3 1986 02:40 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.84 Modelling results at 0245 UT April 3 1986.

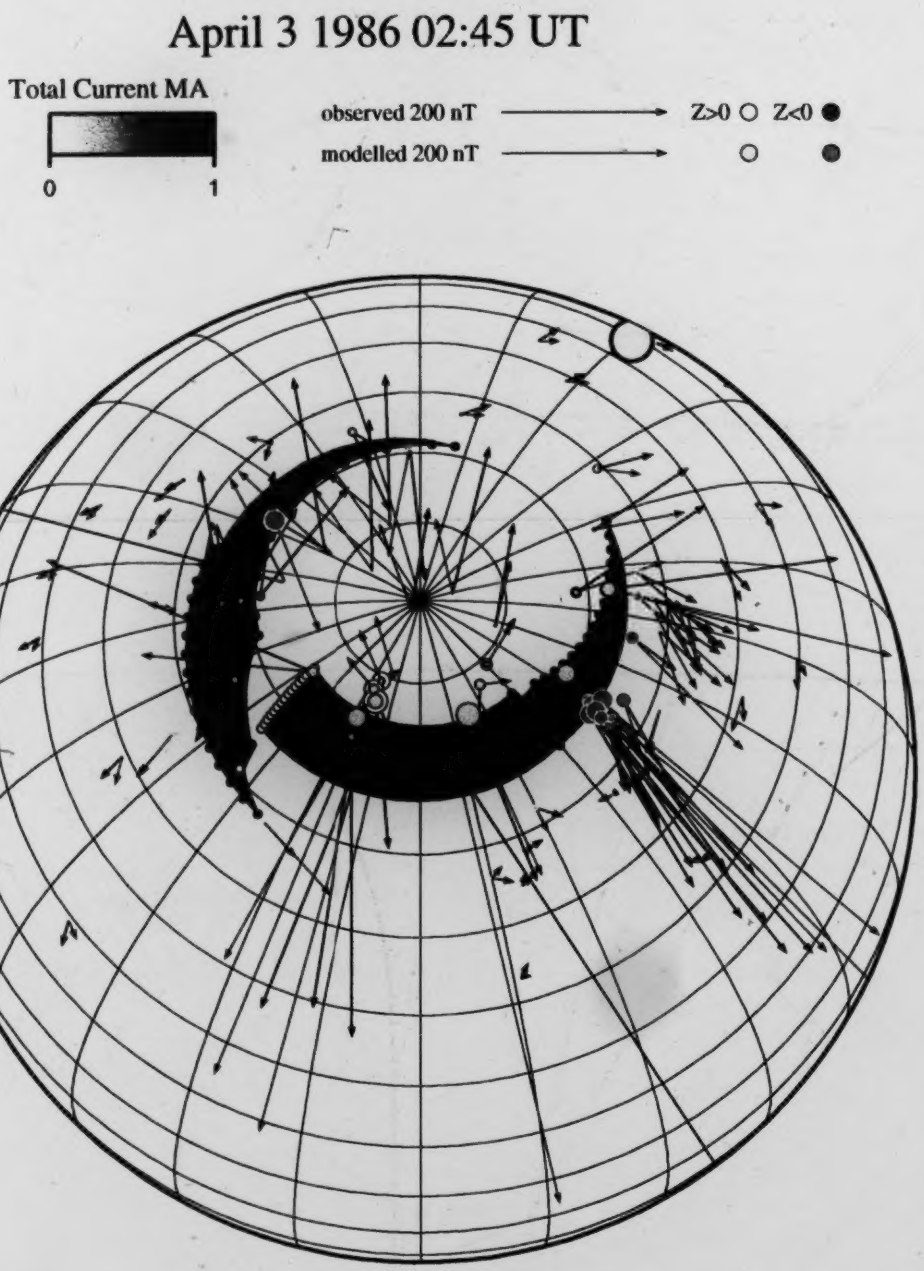
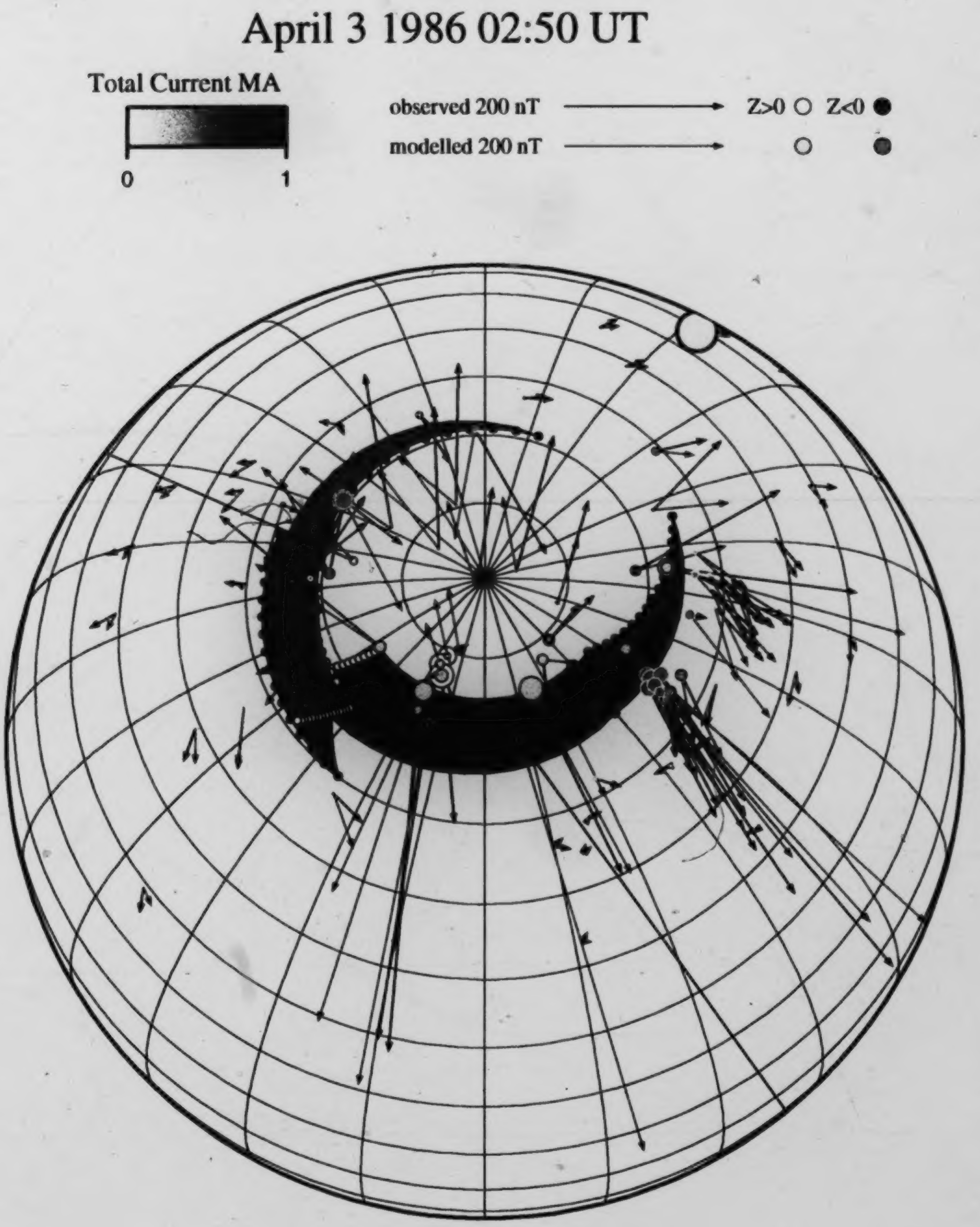
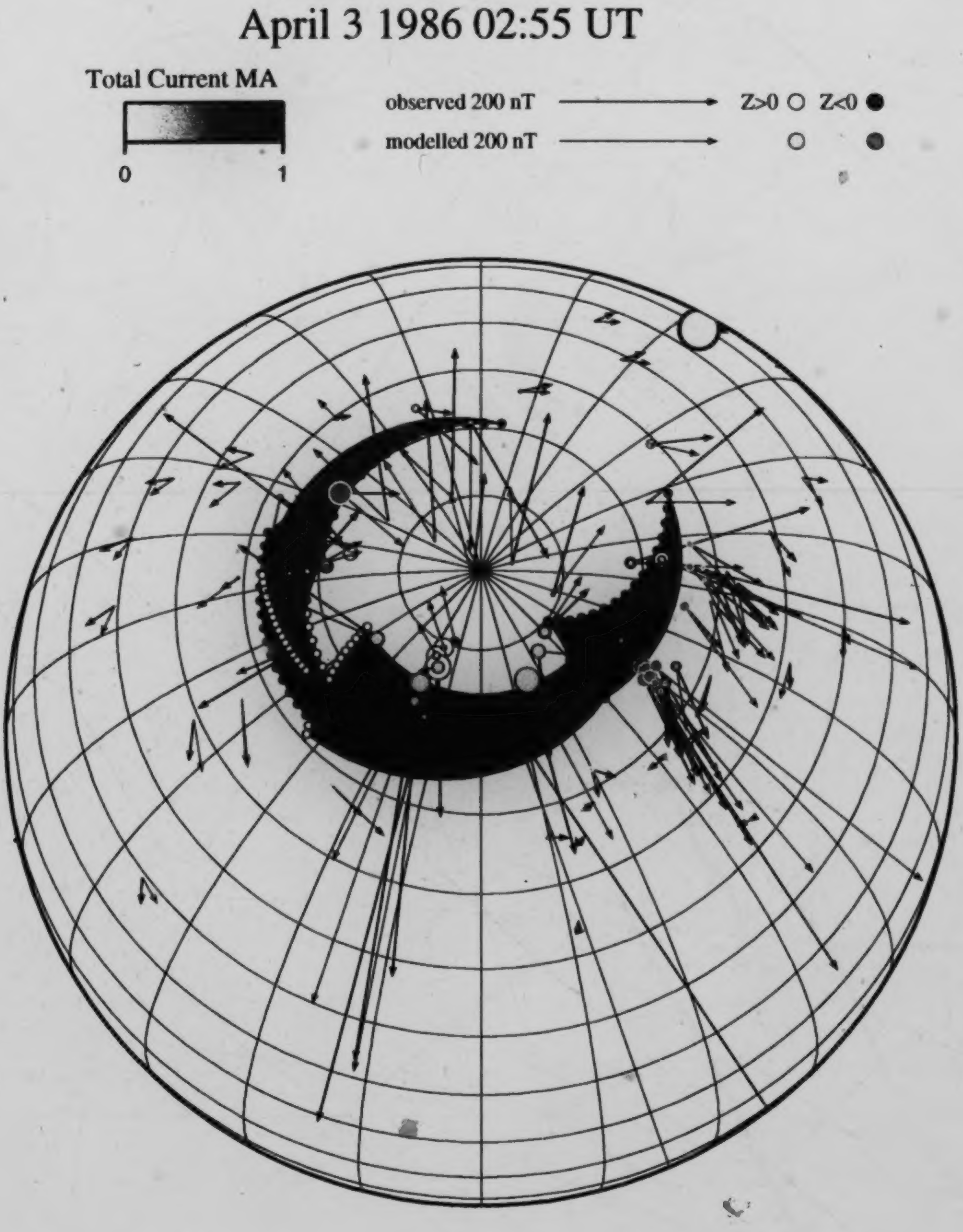

April 3 1986 02:45 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●
○
●

Figure 6.85 Modelling results at 0250 UT April 3 1986.



April 3 1986 02:50 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 6.86 Modelling results at 0255 UT April 3 1986.


April 3 1986 02:55 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0
Z<0

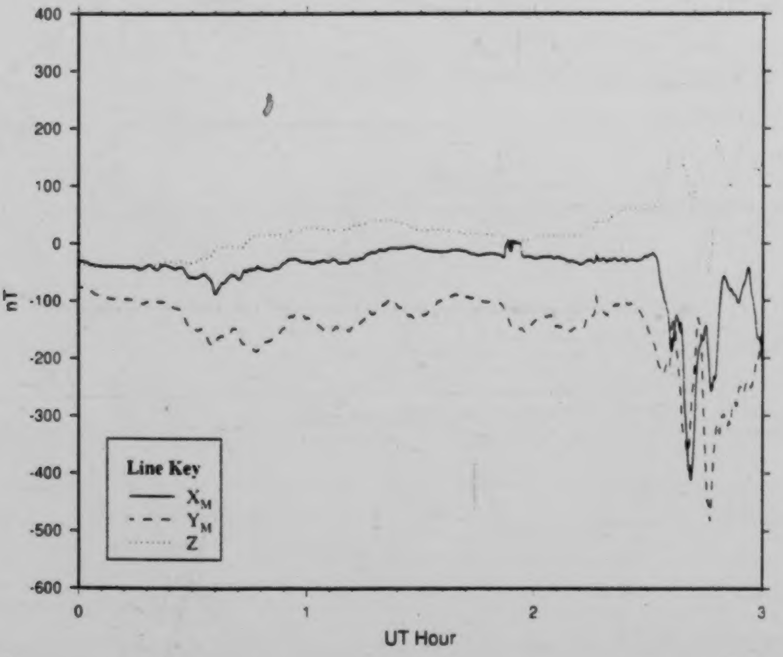
Modelling results (Figures 6.75 to 6.86) show current systems which generally allow a good representation of the data, as may be verified by inspection of the vectors representing observations (black) and those representing model results (shown in red). Not shown is the D_{st}-like parameter used in modelling. It generally had a value of about -25 nT, determining much of the horizontal perturbation at mid-latitudes. For mid-latitude stations not having a pronounced southward component, the effects of auroral oval currents have been to offset those of the D_{st}-like parameter. Referring to Figure 6.74, showing the initial conditions, or Figure 6.75 which does not deviate much from them, one should note that there are basically three current systems present. The morning sector westward electrojet is represented by two joined regions of current flow as before, with a junction at 105° centred dipole magnetic (east) longitude. The evening sector eastward electrojet extends from the early afternoon to late evening sectors. The wedge current region has a large degree of north to south current flow. In many cases it is partly overlain by the evening sector electrojet and is difficult to see. To assist with visualization of the wedge, the reader should bear in mind that the field-aligned currents at the ends of current systems are depicted by 16 uniformly spaced circles. For example, in Figure 6.75, most of the circles are visible at the southern border of the wedge current region (not yet an active wedge) and they are black, indicating downward field-aligned currents. At the northern border, only about half of the (white) circles are visible, so that a region of upward field-aligned current extends considerably farther west than is visible in the figure. The net meridional current flow in the wedge is thus from east to west, corresponding to that expected in the generic view of a wedge, to that expected from midlatitude observations, and to the negative H bays seen on the ground under the wedge electrojets. The south to north currents, however, placed mainly in an attempt to match polar cap perturbations, result in an overall current strength in the wedge which is too large (0.63 MA as compared to 0.61 MA in the morning sector westward electrojet) for this pre-onset time. The net effect of the overlapped eastward electrojet/wedge systems is to place northward field-aligned current in the late evening sector, and to largely cancel the effect of the eastward electrojet there. These results are in accord with observation, and properly interpreted do not give an unphysical result. Figure 6.76 shows a similar configuration for pre-onset conditions. In figure 6.77 the 'wedge' region has been extended further west, and in this western region its current flow is basically south to north. Both midlatitude and polar cap perturbations have been better modelled through such extension. The westward electrojet has extended further from the morning sector into the evening sector, and perturbations everywhere except in the polar cap are well represented. In Figure 6.78, the conditions at onset time (0215 UT April 3) are represented. The eastern end of the wedge current is now well north of the eastward electrojet, allowing those stations having strong southward perturbations to be well represented. Yet more strengthening and northward movement to reduce overlap is clear in Figure 6.79. Figure 6.80 illustrates the same time step but with Region 1/2 currents included. These were colocated with the electrojet (none was placed in the current wedge) and their total current solved for. It may be seen that the match to polar cap perturbations has been dramatically improved by this approach, as have most of the matches to midlatitude perturbations.

The physical interpretation of the final result is also much clearer in this case. The current wedge no longer overlaps the evening sector electrojet and is no longer forced to carry much south to north current flow in order to optimize the polar cap match. The perturbations near the wedge are now clearly associated with its current flow, and the important fact that it trends toward the northwest is clear. The addition of Region 1/2 currents is seen to have brought the model closer to physical reality. Not only has this allowed a reasonable explanation of polar cap perturbations, but by having included their source, the other parameters in the model have been adjusted into a more reasonable configuration as well. The various current strengths now all have a physical meaning, whereas that of the wedge (over 1 MA) did not have a clear meaning when it was showing mainly south to north current flow. The wedge now has 0.55 MA of current flowing northwesterly within it, the westward electrojet remains the strongest system with 0.76 MA, and the evening sector eastward electrojet carries 0.28 MA. In the morning sector there is a total of 1.05 MA of current flowing north to south as Region 1/2 current, and in the evening sector 1.71 MA flowing south to north. Another interesting feature is the intrusion of the westward electrojet into the evening sector. The combined effect of wedge downward and westward electrojet upward currents is to null out eastward component perturbations at midlatitudes in the 0° magnetic meridian (Ottawa and Fredericksburg).

In Figure 6.81, showing the situation at 0230, there is clearly a nonphysical overlap of the eastward electrojet and wedge current systems within the 0° to -20° magnetic longitude sectors. The overall configuration, however, is acceptable, featuring a westward electrojet which has intruded into the evening sector, and evening sector eastward electrojet, and the current wedge north of it and with current trending northwestward. Figure 6.82 shows a model for 0230 which includes Region 1/2 currents. This model does not have any nonphysical features, although the extension of the wedge current into the polar cap, which it shows, appears not to be justified. The morning sector westward electrojet has been strengthened in this model to carry 1.03 MA of current. Some strengthening has undoubtedly occurred, but this number overrepresents that trend. At the northern border of the electrojet, the downward and upward field-aligned currents are very close. Only in the longitude of eastern Greenland (90°) does 1.03 MA actually flow across a meridian. At other longitudes the current flow is reduced and would be roughly comparable to that found at 0230 (i.e. about 0.75 MA). The wedge is slightly weakened from 0230, at 0.49 MA, and the eastward electrojet is slightly stronger at 0.32 MA. The morning sector Region 1/2 currents come out as somewhat strengthened at 1.47 MA, but the evening sector Region 1/2 currents are indicated to have weakened considerably to 0.55 MA. There are several factors to consider in interpreting the results for Region 1/2 currents. First, the total Region 1/2 current has not varied as much as the individual currents appear to have done. The polar cap perturbations essential sum contributions from morning and evening sectors, and if one has come out too small and the other too large, then that sum should not change considerably. Basically, this is to say that the routine may be more sensitive to total Region 1/2 current than to its distribution. Second, the placement of actual wedge current into the evening sector polar cap, in itself likely an incorrect result, would reduce the need for as much contribution to the evening sector perturbations by evening sector Region 1/2 currents. The results, however, have much more obvious physical meaning than those for the corresponding time without inclusion of the Region 1/2 currents. By 0240, the time of Figure 6.83, the wedge was declining and its distinct perturbations were mostly seen at only one station, BLC. Thus the 'wedge' once more had its parameters used to make a current system with large south to north current flow to satisfy goodness of fit in the polar cap. The resulting current flow, at roughly 45° to local meridians, also models quite well the perturbations further south than BLC on the Churchill meridian. It may also be noted that midlatitude perturbations in western North America are modelled with too much eastward component in this case. This is undoubtedly due to the absence of the field-aligned current which must be present to drain the wedge current north of the evening sector electrojet. This may also be said of Figure 6.84. Inspection of the polar cap perturbations, compared to those prior to onset (see Figure 6.75) shows why better convergence has not been possible. The best match to those perturbations could be had by increasing the south to north current in the 'current wedge' at a location near its position in the input model. This is exactly what was done. Moving it northward the nearly 10° which would have been needed to place it where the westward currents really were located (i.e. near BLC) was not within the scope of the program. The program was, however, able to bring the western edge of the westward electrojet further west, attempting to match perturbations at IQL. These perturbations are likely explained by the enhanced northern border currents which have already been shown to have been associated with the westward electrojet. The remaining figures (6.85 and 6.86) illustrate conditions during the decline of the wedge. While the evening sector does not have much to teach in these figures, comparison of the morning/midnight sector results and perturbation shows a much enhanced westward electrojet there. In addition to intruding into the evening sector, the electrojet has widened and increased its total current, the latter from 0.616 MA at 0200 to 0.853 MA at 0255.

During the lifetime of the major current wedge during the 2 to 3 UT period, magnetometer data have been used to derive model current systems which are physically meaningful when subject to proper interpretation. It was shown that models including Region 1/2 currents could be useful at certain times, allowing a clear separation of systems which otherwise overlapped or were deformed from what was used as input (i.e. basically westward current flow in the wedge being redirected to be essentially south to north). Some westward motion of the substorm current wedge took place but the transition bay seen in eastern and central North America (Figure 6.73) is due more to intrusion of the westward electrojet after onset time than to westward motion of the current wedge. The positive eastward perturbations arising from the upward field-aligned current of the westward electrojet, as seen at 0255, remain after 0300, while the wedge itself has by then died away as evidenced by the Fresno data. A close connection between the current wedge and electrojets (in this case the morning sector westward electrojet) is indicated in the sense that the onset of the expansive phase also led to (slower) strengthening of the entire morning sector electrojet. It initially appears that there is a contradiction between the total electrojet current found in global modelling and that found from the Russian meridian chain as illustrated in Figure 6.49. While the overall current from global modelling is 0.853 MA, that indicated for the chain at 0255 is about 0.32 MA. Examination of Figure 6.86 will show, however, that the downward current feeding the electrojet was not completely to the east of the chain. Counting the circles marking downward current, one finds that only five of the sixteen lies east of the chain, so that about $\frac{5}{16} \times 0.853 = 0.267$ MA would be expected to cross the chain in the global model. Comparing to the 0.32 MA indicated by the latitude profile modelling, there is no great discrepancy revealed. It is also interesting to note that the perturbations in Russia immediately south of the auroral oval show very good agreement of the eastward component with the modelling, right throughout the global modelling runs from 0055 to 0255 (Figures 6.28 through 6.34, 6.52 through 6.58, and 6.75 through 6.86). Although early in this period there are some southward perturbations near the subsolar meridian which would indicate that current had already flowed into the ionosphere there, the primary region of downward current for the westward electrojet was over Russia much as these models show.

Figure 6.87 Perturbations at high latitude station Iqaluit from 0 to 3 UT on April 3 1986.



The high latitude station Iqaluit does not show much perturbation until about 0230 UT (see Figure 6.87), when an onset-like signature becomes visible there. In the modelling this corresponds to the time that the westward electrojet penetrated into the high latitude evening sector after onset. Rather large westward perturbations are accounted for in the global modelling by field-aligned current: inspection of Figures 6.80 through

6.85 shows this to be present near the station at the time of the onset-like signature The global modelling generally underrepresents the strength of the horizontal component at the station. Also, Figure 6.87 shows rapid Z component variations at the time the edge of the electrojet crossed over the station. Both of these facts support the idea that not only did the westward electrojet move into the high latitude evening sector at this time, but that its northern border carried an enhanced current. The onset which took place in a localized region of Greenland at 0150 UT produced a signature in the horizontal components at Iqaluit. Pulsation analysis showed an signals in the Pi 1 and Pc 2 bands (50 to 100 mHz and 5 to 15 mHz, respectively) at the time of this onset. These data are not shown but support the contention that the 0150 event was an onset. On the other hand, the onset-like signal at 0230 was not initially accompanied by Pc 1 activity, which supports the contention above that it was not an onset but rather the approach of the active northern border of the westward electrojet.

Referring back to Figures 6.5 and 6.6, it may also be seen that the onset at about 0215 UT showed no strong triggering signal in the IMF. There are changes of about 2 nT in the GSM X and Y components at about onset time but it is not clear that they preceded onset. There is a corresponding Z component change but it is less than 1 nT and many variations of this magnitude can be seen in the data record, but are not associated with onsets. A major change (northward turning) in the IMF took place at only at 0255 but did not appear to trigger an onset.

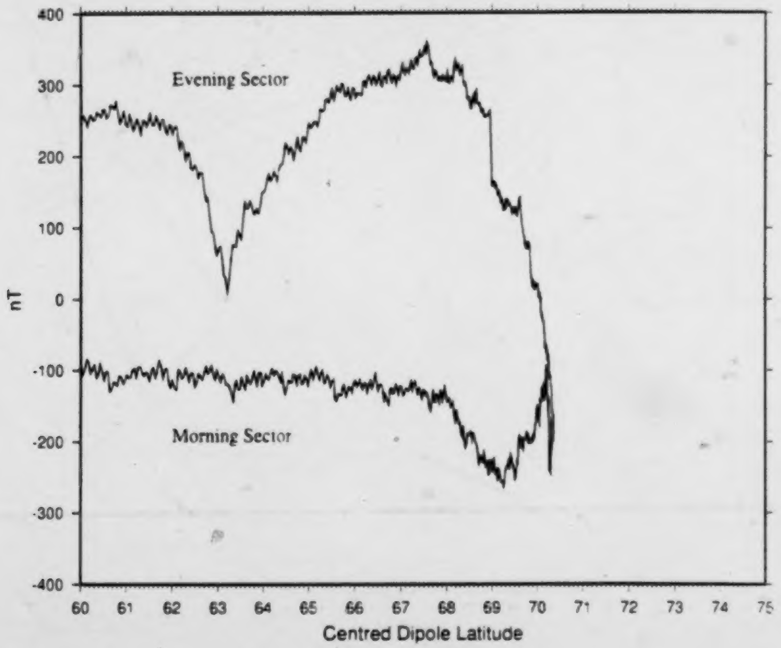
Although mapping is not a major focus of this study, the presence of intensifications which appear to be northward of the main electrojet currents, and others which appear to be in or southward of them, has implications for source regions. In particular, the likely regions which these imply as the sources of expansive phase onsets are quite different. The high-latitude events likely map relatively far out into the tail and in some cases to its flanks. The lower latitude events likely map to the near-Earth tail, a region having received much attention recently in terms of substorm onsets. It appears from the above that both regions may produce similar phenomenology. Elphinstone *et al.* [1995c] have observed a case in which the auroral distribution evolved from a single to a double auroral oval configuration in the sense of having two latitudinally separated regions of bright aurora as observed by Viking on July 28, 1986. In this case also, it appears that there was initially a single auroral oval at relatively high latitude and showing expansive phase activity. Later activity, including a major onset, was primarily at its northern border and eventually the double oval formed, with a very active northern border. This northern border could feature onset-like signatures as it moved over stations. The current wedge which formed later in the evening sector was poleward of the eastward electrojet there (but was not well represented in modelling). It was pointed out that most of this double oval activity took place over the ocean, so that interpretation from ground magnetic data was difficult. Use of Viking images proved invaluable in understanding what took place. We now seek evidence of the enhancement of the poleward border using a different type of data.

h. DMSP passes

Having studied the ground perturbations, which are primarily due to Hall currents, and their relation to other data, attention is now given to low-altitude satellite data from the DMSP F7 satellite. These data may be examined in a consistent fashion among themselves and reference will then be made to their relationship to events deduced above. Global modelling was not done before 0 UT April 3, but positional information about electrojets can be extracted from earlier satellite passes. The position may be deduced from any single pass in which the characteristic signature of entry to and exit from a solenoidal north-south aligned current system may be recognized, without reference to anything else. On April 3 itself, it is of interest to also evaluate the magnitude of the solenoidal currents for several reasons. These include tracking the growth of current systems for comparison to what is seen on the ground, and evaluation of the possible distant effects of the solenoidal system (in the polar cap, for example). A large baseline signal is visible in the DMSP plots as a depression of the morning sector data below that from the midnight sector. This may be seen in Figure 6.88. It aids in identifying the two parts of the passes, but is a systematic data error or source of 'residuals'. To address this problem, the approach here is to use the early UT hours of April 4 as a quiet reference level, much as was done with the ground data as prepared (largely) by Barbara Emery. Since satellite passes on April 3 and April 4 did not occur at the same time, those for the same hemisphere occurring closest in UT were paired. Passes which were close in UT were verified to be similar in coordinates traversed, both in Earth-fixed coordinates (geographic and geomagnetic) and in local time-latitude. Further, the similarity of magnetic perturbations during times of low activity was verified by examining several quiet intervals throughout April 1 to 4, and this supported the use of quiet periods as baseline values for active periods. Due caution must be used to allow for spurious signals due to shadow entry, tape recorder activity, and other possible sources which are not directly due to the magnetospheric currents. The quiet day may or may not contain these same signals: the use of the quiet day primarily allows removal of those residuals of undetermined origin [Langel *et al.*, 1990, p. 56] which repeat systematically but which may be large compared to the perturbations due to currents. Passes are numbered sequentially through the UT day and of course alternate between southern and northern hemisphere passes. The northern hemisphere passes sample the midnight and late morning sectors and the southern hemisphere passes sample the early evening and noon sectors.

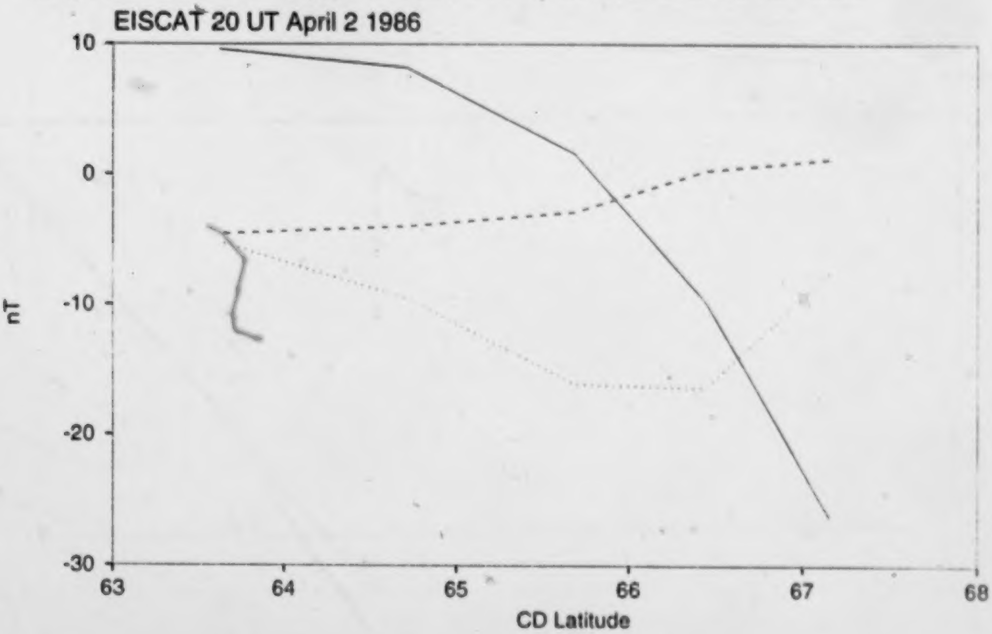
Magnetic field data were transformed from the satellite coordinate system to the local geographic system appropriate to the spacecraft footprint as indicated by interpolation of the ephemeris file, as described in Chapter 3. In practice the geographic Y component indicates the perturbations due to field-aligned currents. In principle it would be better to transform to a local magnetic system and this transformation is not difficult. However, the level of analysis performed here is adequately done with the Y component. Most published DMSP perturbation work appears to have been done in the spacecraft system with a rough comparison to magnetic north-south indicated by an ephemeris plot and what is done here goes a step further than does that approach.

Figure 6.88 DMSP Y component as a function of latitude on April 2 1986 (Pass 24). Traces for subsequent midnight sector passes are shown in the following diagrams.



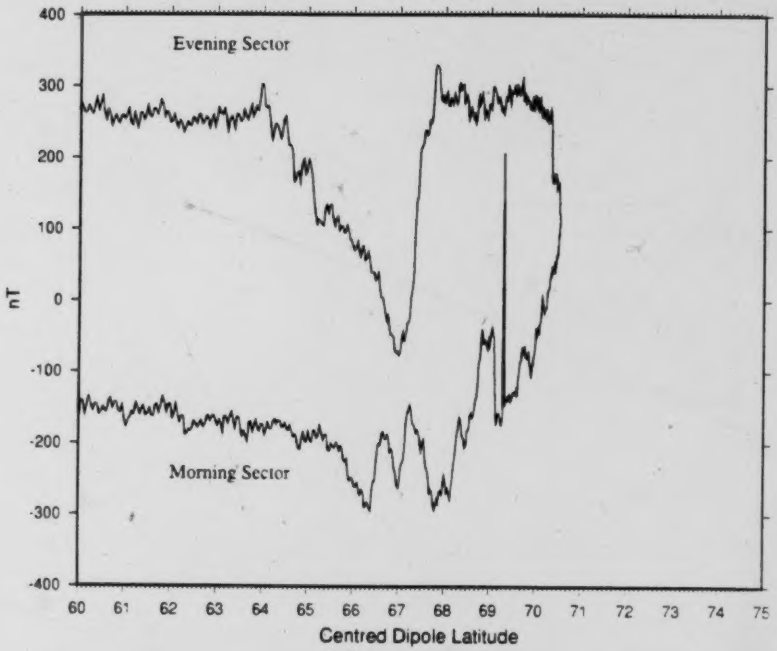
The early development of the westward electrojet through the midnight sector may be examined using DMSP passes from April 2. It has been found above that the electrojet is tilted to the northwest as one follows it westward from Scandinavia. Images and magnetic data suggest that this configuration was maintained at least from about 2100 UT April 2 until substorm onset at 0055 April 3. During that period the midnight sector rotated from east of Scandinavia to eastern Greenland. The images suggest that the tilted configuration was not fixed in magnetic local time but to a certain degree rotated with the Earth. Otherwise puzzling DMSP observations during this period can be explained by assuming that this is what actually occurred. Figure 6.88 illustrates Y component perturbations from DMSP pass 24 during the late UT hours (19.7 to 20.1 UT) of April 2. The Y component has a minimum at 63° which indicates an equatorward Pedersen system of the type associated with westward electrojets (i.e. traversal of roughly balanced Region 1 and Region 2 currents). There may be a second negative signature located north of 69° but the satellite pass does not fully traverse it, suffers from artifacts such as a level shift at 69°, and shows a large background signature. The morning sector (bottom of diagram) has a relatively clear Region 1/2 signature between 68° and 70°. Passes 22 (data not shown; ca. 18.3 UT) and 24 show a Pedersen system apparently fixed in latitude at roughly 63° in the midnight sector. Still earlier, on pass 20 (data not shown; ca 16.8 UT), currents are found in the range 61° to 66° in the evening sector. These currents may reflect remnants of earlier low latitude activity. Throughout the period during which currents at or about 63° were detected by DMSP, inspection of Viking images did not reveal any UV luminosity where it might have been expected from this low-latitude current system. Sulzbacher *et al.* [1980] showed a Triad pass where only the most extreme magnetic perturbations corresponded to regions of auroral luminosity and with no luminosity apparent in the area of a secondary magnetic maximum. In their study, a dense array of stations in Scandinavia allowed the secondary maximum to be correlated with a small enhancement of the northward component as observed on the ground. This enhancement became more pronounced in field continuation to ionospheric altitude. In the case discussed here, the ground observations at the time of the pass, as shown in Figure 6.89, did not show perturbations at the location of the inferred Pedersen current system. The Z minimum at roughly 66° and declining H component suggest that the centre of the westward electrojet was at or north of 67°.

Figure 6.89 Latitude profile at EISCAT meridian for 20 UT April 2 1986. Local magnetic northward component shown solid, eastward dashed, and Z dotted.



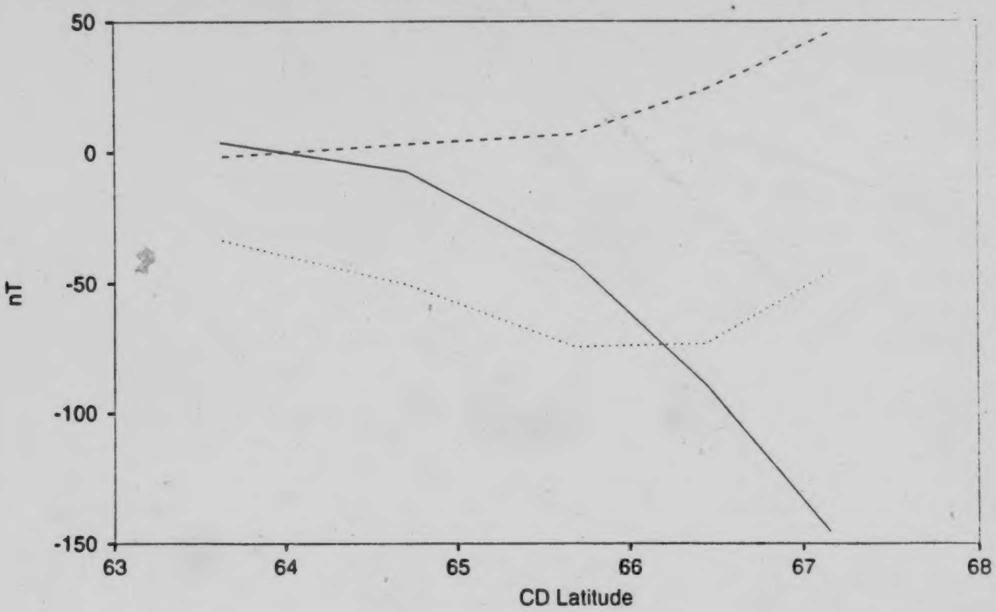
Station coverage is less than in the Sulzbacher *et al.* study but there appears to be little ground effect associated with the Pedersen system detected by DMSP. Ground perturbations are clearly those of a system centred near or north of 67°, but the position of the Z minimum and inspection of records from Bjørnøya suggest that the electrojet is not greatly north of 67°. It is a well-known problem of Scandinavian mainland stations that the entire electrojet often cannot be detected [Untiedt and Baumjohann, 1993]. Pass 24 data, although confused at northern latitudes, do show evidence of a decline, especially as compared to pass 26 (see Figure 6.90). Strong field aligned currents in the evening sector and complex morning sector features were located between 67° and 68°. Thus it appears that the westward electrojet. as evidenced by both data sets, is north of 67°, and that the lower latitude perturbation is not associated with the westward electrojet. This conclusion is supported by comparison with average positions of the electrojets (recall Figure 1.9) which show the 65° to 67° range to be usual for the quiet-time westward electrojet.

Figure 6.90 DMSP Y component as a function of latitude, April 2 1986 (Pass 26).



There appears to be no evidence of field aligned currents that would indicate south to north ionospheric current flow which is normally associated with the eastward electrojet in these DMSP data, whereas statistically it could be expected to be south of the westward electrojet near midnight. The absence of this signature is consistent with the results from the modelling studies of Section 6.b. There, the evening sector eastward electrojet was also deduced to be very weak. Pass 26, (21.4 to 21.8 UT April 2), after a brief southward turning, clearly shows the westward electrojet at higher latitudes, roughly 67° (see Figure 6.90), and no evidence of a current system at 63°. The northern border of the electrojet appears to have become steeper with time, indicating a more concentrated current flow (downward) there. The perturbation has also become stronger indicating more total current flow in the Pedersen system. The Hall system as detected on the ground in Scandinavia, nearly underneath the DMSP pass, had also strengthened as seen in Figure 6.91. There is little evidence of motion of the electrojet, with Z minimum at roughly 66° and declining H component still suggesting that it is at or north of 67°.

Figure 6.91 Latitude profile at EISCAT meridian at 21.6 UT April 2 1986. Local magnetic northward component shown solid, eastward dashed, and Z dotted



Finally, pass 2.28, from 23.15 to 23.5 UT April 2, (Figure 6.92) shows that the current systems have moved northward with little change in overall appearance. It now approaches 70°, the latitude of southern Greenland. The poleward motion is more likely associated with the tilt of the electrojet than with 'growth phase' equatorward expansion of the auroral oval, which would cause the opposite effect. That the electrojet is at roughly 70° is supported through inspection of the EISCAT latitude profile shown in Figure 6.93, with a Z minimum at roughly 66° and declining H component toward the north. This conclusion is backed up by a positive Z perturbation seen in Bjørnøya (data not shown), north of the stations shown in the profile. In west Greenland at this time modelling indicated that there was an electrojet was stably centred at about 75°, so that once more DMSP appeared to be showing field-aligned currents well south of the electrojet. The total extension northward of the westward electrojet appears to have been more than 5° and the electrojet appears to have largely rotated with the Earth during this time period. In terms of local time it appears as if the electrojet was a tilted structure moving eastward at a rate of about 1 hour of local time per hour. This interpretation explains the DMSP observations at constant local time, ground observations at EISCAT and in Greenland, and is consistent with Viking images. The consistent finding that DMSP detections of field-aligned current signatures were south of where ground perturbations and modelling indicated the electrojet to be is also consistent with the idea that there was a double structure to the auroral oval.

Figure 6.92 DMSP Y component as a function of latitude on April 2 1986 (Pass 28). No clear morning sector features are seen.

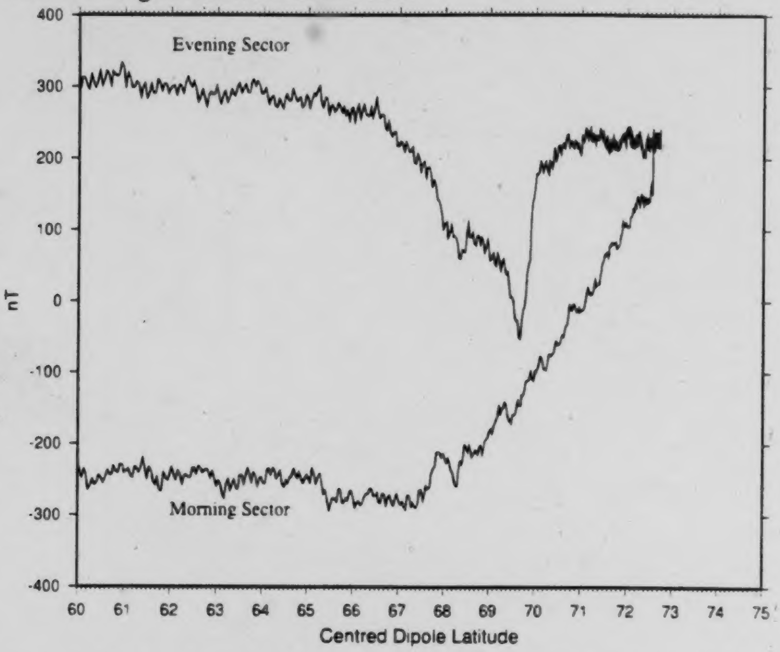
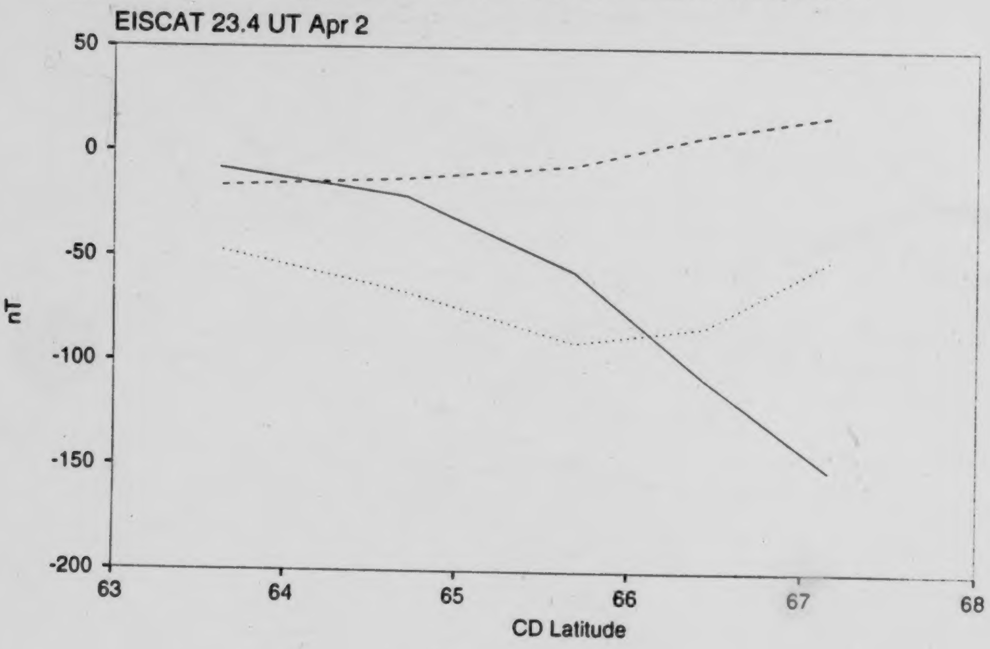
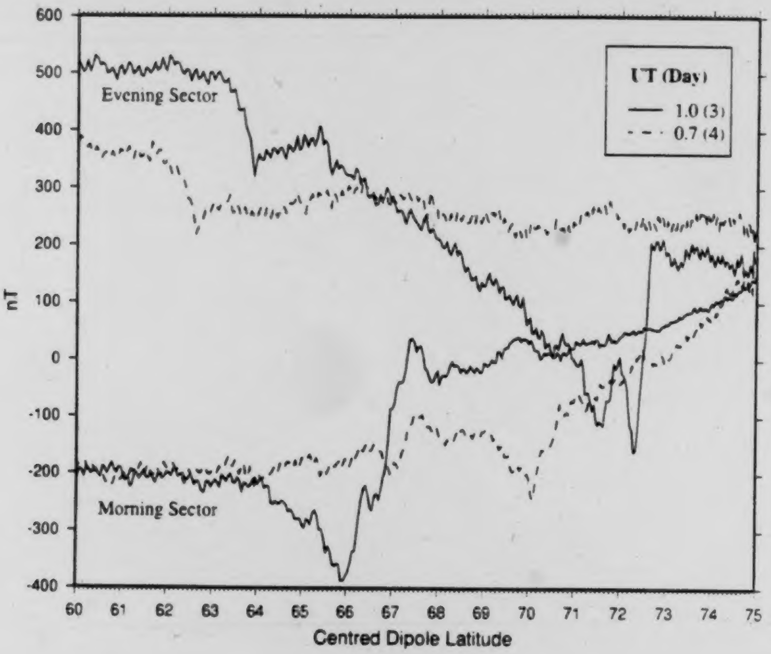


Figure 6.93 Latitude profile at EISCAT meridian at 23.4 UT April 2 1986. Local magnetic northward component shown solid, eastward dashed, and Z dotted



The next northern hemisphere DMSP pass was pass 2 on April 3 from 0.83 UT to 1.25 UT. Traversal through the midnight sector occurred shortly after substorm onset. The perturbations shown in Figure 6.94 include those from approximately the same time on the April 4 quiet day. Several changes may be seen by comparison with the previous pass (pass 2.28, see Figure 6.92). The midnight sector trace (top) descends smoothly from about 500 nT at about 65° to -100 nT with a northern border poleward of 72°. This more northerly location of the electrojet northern border is consistent with its expected location due to the tilt discussed above. The border appears to be doubled: south of the narrow downward sheet at 72.5° there is a narrow sheet of upward current from 72° to about 72.2° and a sheet of downward current equatorward of 72°. Viking images from approximately the same time also clearly show a double structure (see for example Figure 6.36) although the northern border appears more irregular than the DMSP pass might indicate. DMSP indicates a broad region of upward field aligned current equatorward of this, extending nearly 10° to about 63°. This region is much broader than the southern region of upward flow of pass 2.28, which extended only about 3°. The morning sector (lower) part of the trace for April 3 is smooth between 67° and the northern limit of pass, with downward current at 66° to 67° and lesser upward current from 63.5° to 66°. The main morning sector feature on the April 4 pass may be interpreted as the contracted auroral oval at about 70°.

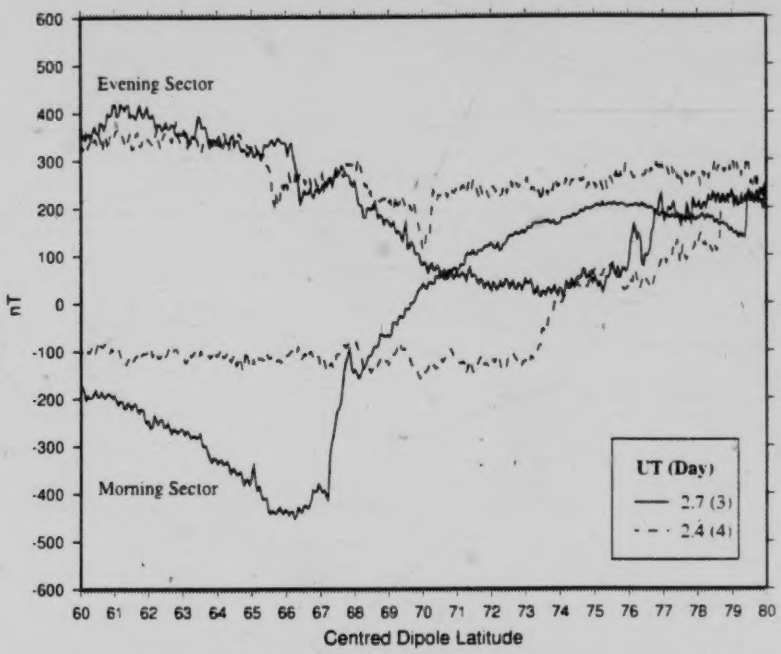
Figure 6.94 DMSP Y component for Pass 2 April 3 and Pass 1 April 4, 1986. April 3 trace is solid. April 4 trace dashed. Dips at 62° and 64° are artifacts.



In the midnight sector, there appears to be a level shift between the northern and southern electrojet borders. This indicates net upward field-aligned current, which is to be expected as modelling indicates that this pass is near the western edge of the substorm current wedge. The concentrated current at the northern border is about as strong as in the previous pass, but the diffuse current to the south is stronger. Away from the midnight sector, the morning (about 6.5 MLT in the auroral oval) currents also appear unbalanced with net downward current. Of significance to the claim that the westward electrojet is tilted northward is the fact that the clear morning sector perturbations are found at a latitude 6° south of the midnight sector perturbations.

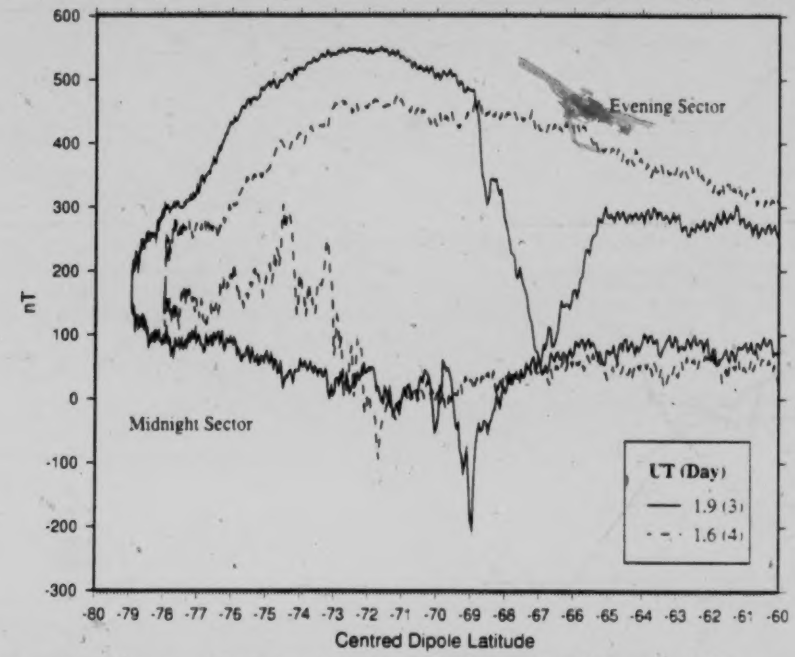
It is of interest to calculate the field aligned currents detected *in situ* by DMSP, using the infinite sheet approximation. Recalling from Section 4.3 that if a Cartesian coordinate system is chosen with z vertical, y eastward, and x northward, the general formula relating an eastward perturbation ΔB_Y to the current density J_z causing it is $J_Z = \frac{1}{\mu_0} \frac{\partial}{\partial x} \Delta B_Y$, the current density in units of A/m^2 may be derived. To study the gross effects of the current systems as opposed to the detailed distribution of the current density, the zones of perturbation will be approximated by a linear change in the field with latitude. In this approximation, the derivative with latitude (along x) is constant and equal in value to $\frac{\Delta B_Y}{\Delta x}$. The linear current density, here denoted i_z, is simply the integral of the current density along x, or $i_Z = \frac{\Delta B_Y}{\mu_0} = 8 \times 10^{-4} \Delta B_Y$ in A/m if B_Y is given in nT. The total current is the product of this linear current density by the east-west extent of the system, in the approximation that the linear current density is constant (single spacecraft observations cannot show to what extent this is instantaneously true). For pass 2 on April 3, just after substorm onset, the midnight sector perturbations (see Figure 6.94) comprise a 660 nT perturbation due to upward current from 63° to 72°, or over 1000 km. This corresponds to a current density of .5 μA/m^2 or a linear current density of .5 A/m. The current density is lower by about a factor of two than that given as typical of active periods by Iijima and Potemra [1978] and the perturbations appear to be about half as large as those of their example plots near the time of substorm onset. The northern border in their study shows a jump of 360 nT in about .1° or 11 km, corresponding to a much higher current density of 25 μA/m^2 or .28 A/m. It is perhaps more meaningful to compare the linear current densities to those given by Fujii *et al.* [1993] based on passes of Dynamics Explorer (DE) 2 which were located with respect to the auroral patterns observed simultaneously by DE 1's imager. The actual linear current densities found here are similar to or larger than those given as typical of the surge and middle surge regions as defined in that study, and the net upward current, here .23 A/m, is very similar to what they found. DMSP results presented here are thus not unusual in linear current although the degree of current density concentration seems to vary considerably. In the morning sector part of this pass, there is a downward current located between 66° and 67° and a smaller upward current between 64° and 66°. The linear current densities of .32 A/m downward, and .16 A/m upward, are similar to those found by Fujii *et al.*, although their difference, a .16 A/m net downward current, is slightly larger than 0.09 A/m found by those authors in the east of bulge region

Figure 6.95 DMSP Y component for Pass 4 April 3 and Pass 3 April 4. 1986 Northern plot limit has been extended northward as compared to the previous plot



The next April 3 DMSP northern hemisphere pass (4) takes place between 2.5 and 2.9 UT. after the second major substorm onset but east of the region of its most intense magnetic perturbations. The perturbations observed are shown in Figure 6.14 and those in the midnight sector significantly differ from those of pass 2. Upward current starts to be found at roughly the same latitude (66°) on the equatorward side, but no sharp northern border is found. The distribution of currents extends considerably farther north and shows some complex structure there. This is quite consistent with the finding from modelling that a more northerly current system had become active about 0130 and remained active. As compared to the April 4 quiet day, there is some evidence of net upward current. However. the currents are in general of lesser magnitude, with some 300 nT, or .25 A/m of overall perturbation, and only about 0.04 A/m of net current. In the morning sector, there is a smooth ramping down of the Y perturbation from the polar cap to about 68°. As in pass 2, there is a distinct downward current at this latitude, but the more southerly region of upward current now appears to extend to roughly 59°. These two northern hemisphere passes sample the midnight and postdawn sectors, but with the strengthening of the evening sector eastward electrojet (as evidenced by the modelling and also by the growth of the AU index as seen in Figure 6.7), it is necessary to consider also the possible effect of field-aligned currents in the evening sector.

Figure 6.96 DMSP Y component for Pass 3 April 3 and Pass 2 April 4, 1986. In this southern hemisphere pass, the evening sector is at the top, midnight at bottom



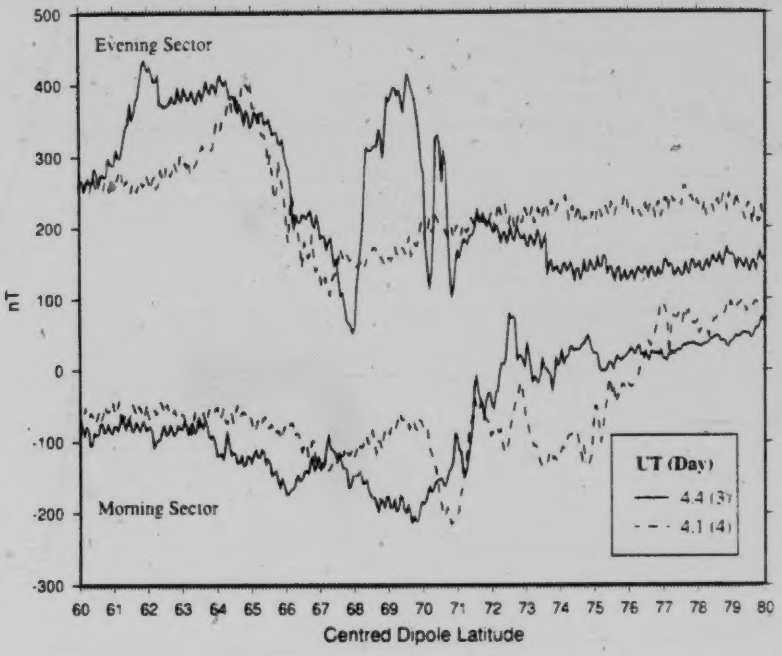
Conjugate data from southern hemisphere pass 3, from 1.7 UT to 2.1 UT, are shown in Figure 6.96. It gives some indication of the strength of the evening sector currents, although they likely increase after the pass. The evening sector perturbation is clearly defined, with a 400 nT drop at the poleward border from -67° to -69°, and a 300 nT recovery between -67° and -65°. The poleward current is .32 A/m upward and the equatorward .24 A/m, leaving a net upward current flow of .08 A/m. Although the linear current density magnitudes are close to those found by Fujii *et al.* [1993], they found current flow in a comparable MLT sector was balanced (i.e. no net flow). It may be noted that the actual current imbalance noted here is at the northern border and that there appears to be a separation of the symmetric main part of the perturbation from the northern border. It is possible that more than one current system is being sampled, with the southern part being similar to the system crossed by DE 2 in the cited study. In any case, for the present purpose, it is noted that a linear current density of at least .24 A/m must be attributed to the FAC pair in this region.

The study of field aligned currents using a low Earth orbiting satellite (DMSP) has allowed changes in the latitude of field-aligned currents to be traced. A more important result is the confirmation of what the modelling appeared to indicate: that at certain times there is an active poleward border of the electrojets. In addition to coming from the modelling through the presence of a strong current system at the poleward border, this is manifest through onset-like activity as observed from the ground, or from poleward motion over stations. From orbit, the corresponding signature is that of a very concentrated field aligned current along the electrojet border.

i. Recovery

The decrease in magnitude of the AL index (refer back to Figure 6.7) after 0300 indicates that westward electrojets, either associated with expansive phase activity or with the driven system, decreased in intensity as would be expected after a northward turning of the IMF. There appears to have been no onset caused by this decisive northward turning. The AU index increased at approximately 0300 and then decreased. By 0400 both indices had decreased significantly, and values approaching those before the southward turning were approached by about 0500. In this regard it is useful to examine the midnight sector perturbations from DMSP at roughly 4.4 UT, shown in Figure 6.97. The main perturbation appears similar in form to that seen in previous passes, with distinct northern border, only now shifted southward to about 68°. Clearly separated poleward from it is at least one up-down current pair. According to Iijima and Potemra [1978], this is a typical recovery phase signature. The detachment of a poleward current system may be part of the development of the poleward portion of the double auroral oval. On the ground, strong activity is observed to continue at Bjørnøya until roughly 0500: this is consistent with the persistent northern border current system seen by DMSP. This is a further argument that the 'onset' seen at about 0130 at Bjørnøya was associated with a different current system than the main westward electrojet, and apparently with an enhanced northern border system.

Figure 6.97 DMSP Y component for Pass 6 April 3 and Pass 5 April 4, 1986.



j. Later activity

After the aforementioned 'recovery' phase, the period between 0600 and 0800 features a return to disturbed conditions. IMP-8 IMF values are not available for this time period, and GOES geosynchronous data are fragmentary. The early part of the period can, however, be studied with Viking imagery and ground magnetic data.

Figure 6.98 Viking auroral image (Camera 0 LBH emission) at pre-onset time 0607 UT April 3 1986. Midnight meridian is shown. Bright spots at right are star images.

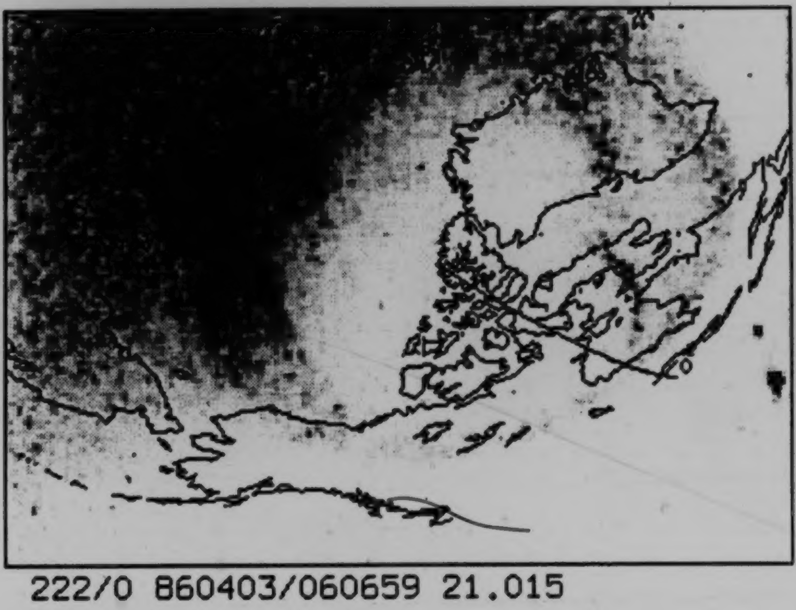


Figure 6.98 shows an overview of the auroral oval as seen from Viking before the 0600 UT decrease in the AL index. The oval appears generally very wide in north-south extent with patchy luminosity in the postnoon sector and in the early morning sector. Only in the evening sector does the oval not appear wide. Figures 6.99 and 6.100 show images from immediately before and immediately after an expansive phase onset. These images permit the onset to be localized in the Churchill sector and almost exactly magnetic midnight. The onset time may be specified to be very shortly before 0623 UT from these images.

Figure 6.99 Viking camera 1 image immediately before expansive phase onset Image has been contrast-enhanced.

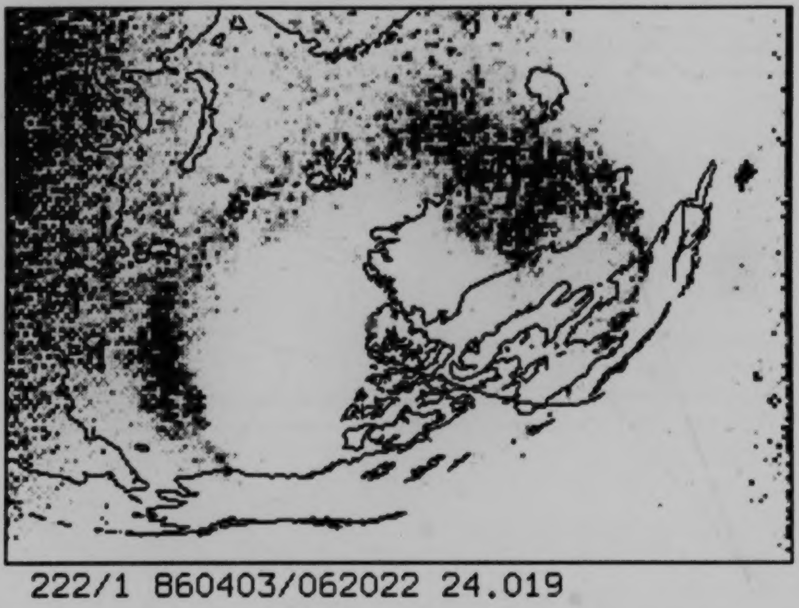

222/1 860403/062022 24.019

Figure 6.100 Viking camera 1 image immediately after expansive phase onset in the vicinity of Churchill. Image has been contrast enhanced in exactly the same manner as Figure 6.99.

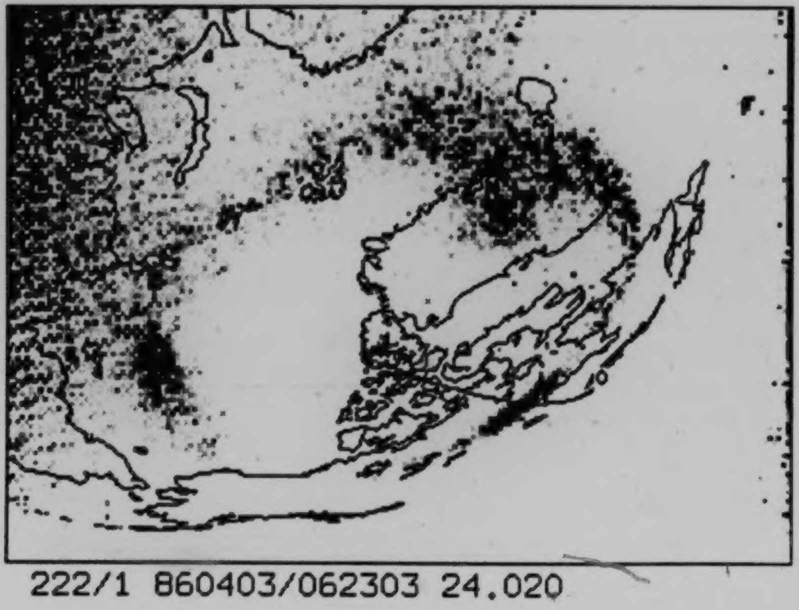
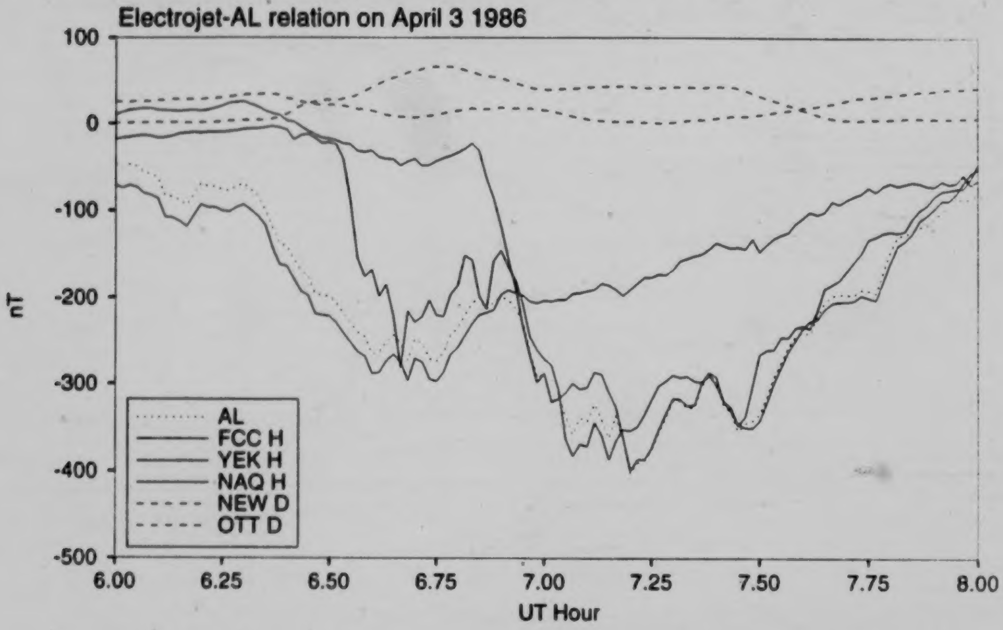
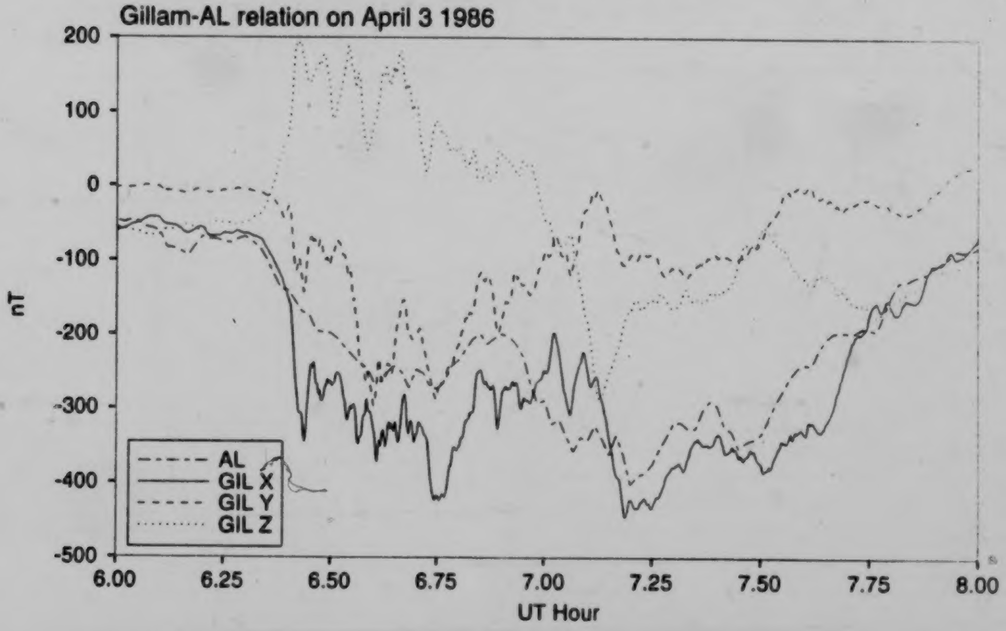

222/1 860403/062303 24.020

Figure 6.101 AL index and magnetic perturbations 0600-0800 April 3 1986. AL index (dotted) is dominated by NAQ (solid trace closely following AL before 7.0 UT) before and after 0623 onset. FCC (solid with drop at 6.55 UT) and YEK (solid with drop at 6.9 UT) show onset-like behaviour only well after onset time. Midlatitude D perturbations at Ottawa and Newport are shown by dashed lines. AL index and NAQ H component are not identical due to differing zero points.



With an onset time of 0623 UT in mind, the details of the AL index as shown in Figure 6.101 may be examined. Although the onset occurs in the Churchill (FCC) sector, the AE station FCC is not contributing to the AL index at onset time. Perturbations at FCC at 0623 are very small and it is in fact NAQ in Greenland at roughly 0300 MLT which dominates the index. This station shows a decreasing trend which started several minutes before onset at approximately 0617. Further west, Yellowknife (YEK) starts to show negative H at onset time, but does not have a major expansive phase decrease until about 0650, at which time both it and FCC decrease and start to closely follow the AL index. The D component in eastern North America decreases at onset time while in the west it increases. Both changes are of about one hour duration. The actual station closest to which the onset occurs is Gillam whose magnetic data is shown in Figure 6.102. Activity is seen to be south of Gillam from the large positive Z component. The AL index is shown for comparison and it is seen that the H component at Gillam is considerably below the AL index during much of the period of activity. A 75-station AE index computed by Kamide [B. Emery, personal communication, 1995] has values of about 400 in a double peak (not shown) at about this time and this is about what Gillam indicates.

Figure 6.102 AL index and magnetic perturbations at Gillam April 3 1986. Large changes in Gillam X and Z at onset time indicate onset in this sector. X and Y at Gillam are essentially identical to centred dipole X_M and Y_M.

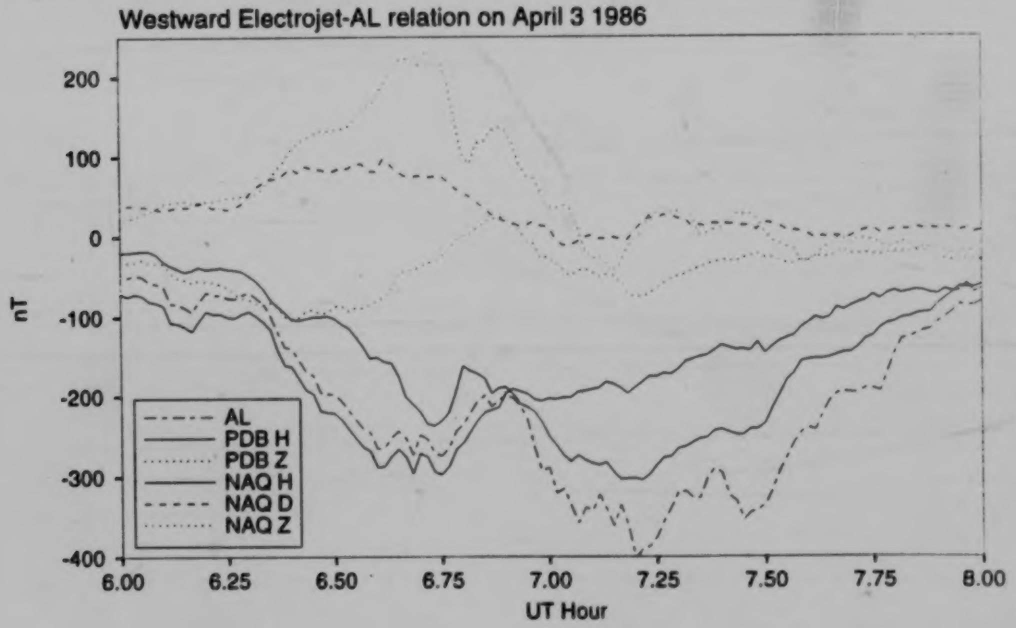


Particularly during the early part of the active period, when the 12-station standard AL index in fact reflects westward electrojet activity in Greenland, the expansive phase currents are not well represented by the index and exceed it considerably in magnitude, at times by more than a factor of two. In this case the expansive phase activity is considerably south of the nearest index station (Churchill). This is the opposite of the situation before 0 UT, when the index did not show the expansive phase activity in Greenland which was north of the Narssarssuaq AE station. In both cases, of course, the index does not accurately show activity at latitudes not close to those of the AE stations. This effect is already known, as indicated in the discussion of the auroral electrojet indices in Section 3.a.1.b, but the examples given here illustrate that even activity which is relatively close to the AE stations can be missed.

While illustrating that the AL index can be quite misleading even in a case where the activity is close in local time to an AE station, this example also illustrates that good station coverage in the form of north-south chains is important to a more sophisticated modelling effort, and the importance of the Z component. The Z component at Churchill (data not shown) in fact quickly goes strongly positive at the time of substorm onset at Gillam. It is possible that a modelling routine sensitive to the Z component and its rapid variations (such as Automated Forward Modelling) could be used successfully in the absence of the station at Gillam. It is difficult to conceive that a routine not adequately incorporating the Z component could model this onset.

Another reason to examine the magnetic signatures associated with this onset is to see whether features of the 0055 onset are seen, in particular electrojet motions. Figure 6.103 shows magnetic data which may in part answer the question. Attention is first drawn to the H and Z components at Poste-de-la-Baleine (PDB). The negative H perturbation is smaller than that at Narssarssuaq and the Z component is essentially a mirror image of H. This indicates that the electrojet is well south of PDB and the constancy of the Z/H ratio shows that the electrojet is essentially stationary near PDB. At Narssarssuaq, a small negative Z component indicates that the station is relatively near, but somewhat south of, electrojet centre at onset time. While H increases in magnitude, Z does not appreciably. This indicates that the centre is coming closer to the station, that is to say moving southward. This tendency continues throughout the initial activity (from 6.25 to about 6.75 UT) with the Z component eventually becoming positive. Close inspection of Figures 6.101 and 6.102 will show some similar evidence of southward motion also. These observations support the modelling conclusions that electrojet latitudinal motion can occur at onset.

Figure 6.103 AL index and magnetic perturbations at PDB and NAQ, April 3 1986. The constancy of Z/H ratio at PDB until about 6.75 UT (0645) indicates that the electrojet over this station did not move, while the increase in Z at NAQ from onset time at 6.25 to 6.75 shows evidence of electrojet southward motion. NAQ H component is a solid line (compare Figure 6.28) closely following AL (dot-dashed) before 0700, while PDB H is above AL until that time. Dotted line nearest top is PDB Z and central dotted line is NAQ Z. Dashed line is NAQ D component.



k. Summary

For this event, the general context of the observations has been discussed and the magnetic perturbations have been modelled on a global scale. The following events and configurations are noted as associated with the current systems:

1) active evening sector at high latitude (75°) with onsets at the western extremity of the westward electrojet

2) eastward electrojet in evening sector is at all times weaker than westward electrojet

3) southward motion of eastward electrojet preceded an onset (0055)

4) morning sector electrojet had southward motion of its southern border at midnight sector onset (0055) but not at evening sector onset (0215)

5) midnight sector onset was equatorward of previously existing westward electrojet

6) evening sector eastward electrojet had significant tilt

7) evening sector current wedge had significant tilt (0215)

8) high latitude morning sector onset occurred at poleward border of electrojet (0125)

9) absence of clear solar wind triggering for any onset

10) enhanced poleward border current systems after onsets

Many of these behaviours are not recognized in the 'canonical' model of substorms but an event rather similar to the 0055 UT onset has been discussed in the context of the 'double auroral oval' by Elphinstone and colleagues [1995a, 1995b]. Results presented here indicate that the northernmost 'oval' of such a system may in fact be the extension of the morning sector westward electrojet. It appears that the northern border of the westward electrojet in the morning sector may also be associated with the double oval. In the cited study, activations at both electrojet borders are discussed. This happened also during the period discussed in this chapter.

From a technical point of view, the routine proved successful in following a sequence of global events provided that a suitable initial model was found. Such a model can be, and was, arrived at by sequential attempts at modelling. The global configuration could be meaningfully followed for at least an hour based on such a model. Addition of current systems associated with Region 1/2 currents improved the fit in the polar cap and at subauroral latitudes in global modelling. Latitude profile modelling was shown to give a meaningful, automated way of determining electrojet parameters.

Chapter 7. CDAW 9A - April 1, 1986

A prominent substorm onset occurring at approximately 18:50 UT on April 1, 1986 was used as prototypical of the "eye" appearance of UV manifestation of substorm expansive phases as seen from the Viking satellite [Rostoker *et al.*, 1987]. It was also discussed with focus on use of satellite optical data to interpret substorm onsets by Murphree and colleagues [1991]. Both of these papers present high quality Viking imaging of this isolated substorm. However, study of this event was initially not undertaken in this project, due to the poor quality and coverage of ground magnetic data in the CDAW CD-ROM database. A larger data set put together by Ahn and coworkers [1995] allowed them to study this event using the KRM method, with 'instantaneous' conductivity based on Viking images. An inversion method based solely on midlatitude variations also permitted Sergeev and Vagina to study this event in collaboration with workers from the Viking project [Sergeev *et al.*, 1994]. The availability of these studies, the generally favorable placement of the substorm current wedge relative to observatories, the indications that this is a relatively simple and isolated event, and especially the kind furnishing of adequate data and AMIE results by Barbara Emery [personal communication, 1995], finally led to the inclusion of this event in the present study. A more recent study based primarily on AMIE and GOES results [Lu *et al.*, 1997] became available after the work reported here was complete. The initial onset of at least two (see Table 7.1) has received the attention of other workers. Here also the emphasis is on, and detailed modelling will address, this 18:50 onset, through use of global data from 18:00 to 19:30 as in the Ahn *et al.* study. Viking imaging is in any case not available after about that time.

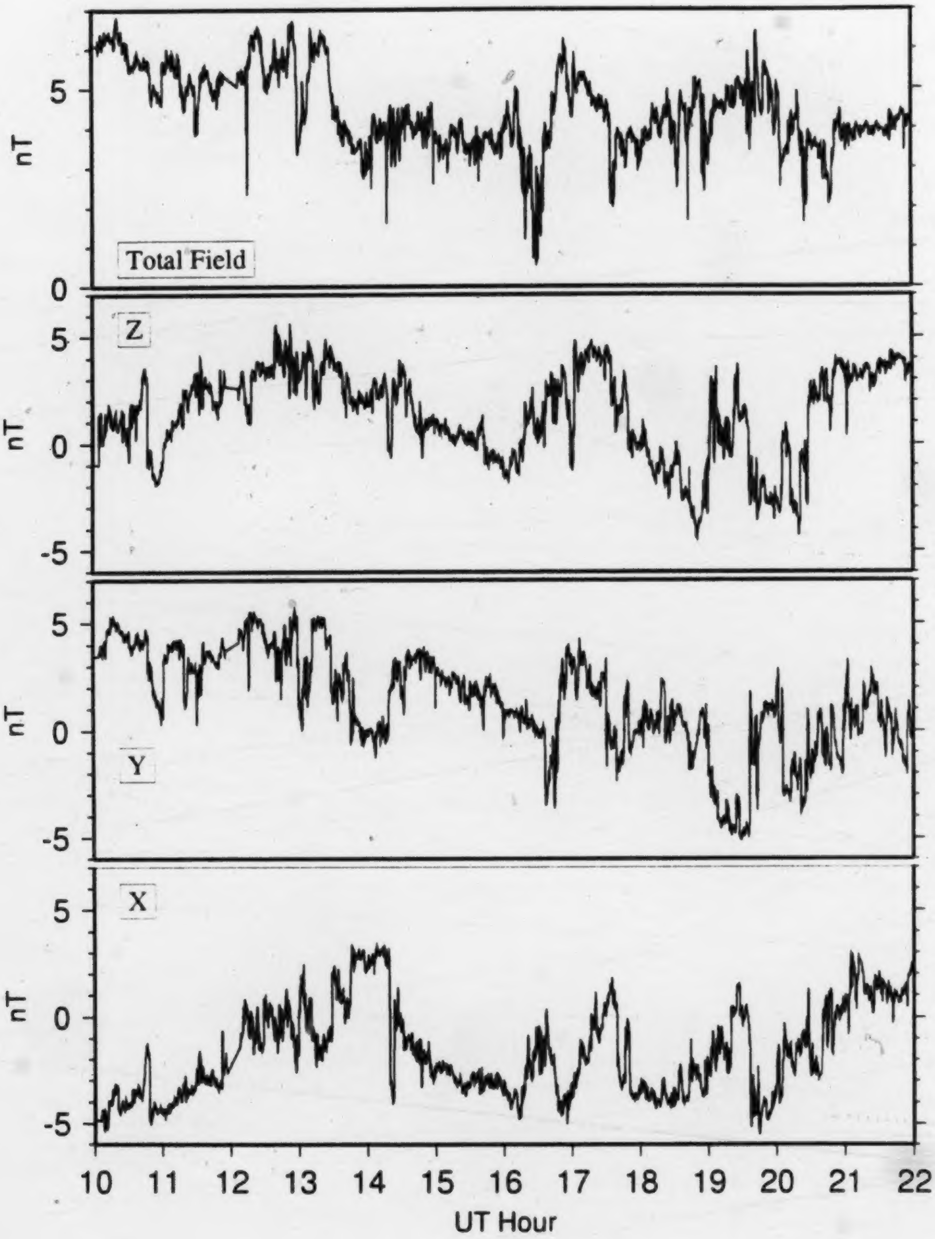
a. Chronological Summary

Here the interval studied in detail is 1830 to 1930 UT on April 1 1986, which features a major onset. The main events relevant to the discussion, and sources of inference, are summarized in Table 7.1. Within the period to be modelled, 1800 to 1930 UT, the substorm sequence is relatively simple. After a period of southward IMF (which began at 1750), an onset occurred, apparently triggered by northward turning. Figure 7.1 shows the IMF as seen at the IMP-8 spacecraft on much of April 1 1986. The spacecraft was located roughly 30 R_e sunward of Earth and 21 (10 UT) to 26 (22 UT) R_e duskward of the Earth-Sun line. The IMF was northward throughout most of the period shown, with brief southward excursions at 1100, about 1600, 1800-1900, and 1930-2030. The variability seen here is usually associated with Alfvénic fluctuations in the solar wind, whereas the more extended periods of southward field seen in the previous chapter are likely associated with larger scale structures ("clouds"). This variability also features impulsive southward excursions of the IMF, seen at 1425, 1450, and 1700.

Table 7.1 CDAW 9A Onset on April 1 1986

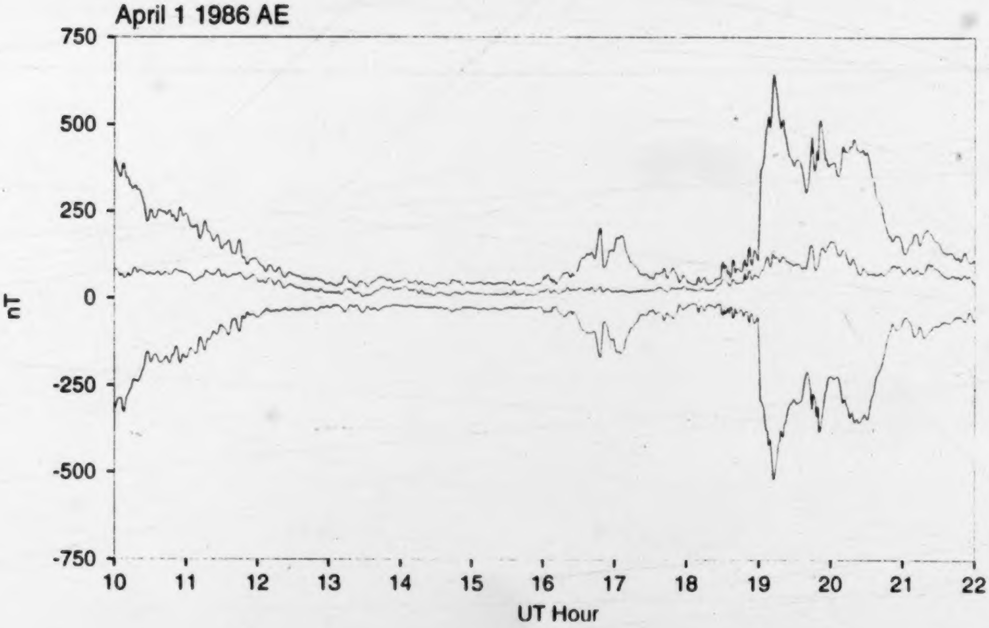
Time (UT)	Activity	Observational Basis
1000-1200	decline from earlier active period	AE index
1425	impulsive change in solar wind including southward B_Z pulse	IMP-8
1445	brief, localized auroral brightening in evening sector	Viking
1750	southward turning of IMF starts	IMP-8
1810	growth of evening sector eastward electrojet	Greenland ground magnetics
1830	eastward electrojet has doubled in strength but not moved	Greenland ground magnetics
1850	start of IMF northward turning onset in midnight sector rapid poleward motion of northern boundary of westward electrojet	IMP-8 Russian stations; Viking; electron injection Russian stations
1910	20-fold increase in westward electrojet	Russian stations (primarily)
1930	eastward electrojet has moved south 3 degrees; westward electrojet declining	ground magnetics
1935	southward turning of IMF	IMP-8
1945	second major onset	AE; electron injection
2030	definitive northward turning of IMF; rapid decline of activity	IMP-8; AE index

Figure 7.1 IMP-8 GSM magnetic field on April 1 1986. From top, the total field, Z, Y, and X components are shown.



The auroral electrojet indices seen in Figure 7.2 illustrate a general correlation between periods of southward field and magnetic activity. The 1000-1200 period, during which a general decline of activity seems to occur, is not discussed here except to note that the southward turning at 1100 appears to have little magnetic effect. After this early interval, the long periods during which the IMF was northward appear to be magnetically quiet. The brief southward turning at 1600 is followed by about an hour of enhanced activity. The period of southward IMF from 1800 to 1900 shows impulsive activity in both the AL and AU indices, superposed on a slowly rising trend. The northward turning at about 1900 appears to correlate with a major AE increase, and AE remains elevated through periods of alternating northward and southward field which follow, until it starts to decline after a definitive northward turning at about 2030.

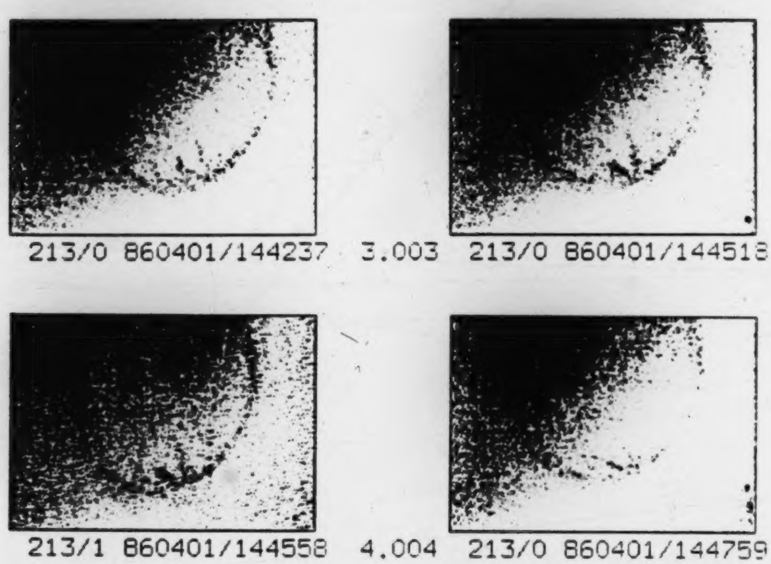
Figure 7.2 Auroral electrojet indices on April 1 1986. Distinct increase at about 1900 UT corresponds to a major onset. AE at top, AU in middle, AL at bottom.



Having noted the impulsive southward turnings and the general associations of southward turning with magnetic activity, it is interesting to examine Figure 7.3 which shows an impulsive brightening of the poleward border of the auroral oval, seemingly associated with evening sector sun-aligned arcs which are a common feature when the IMF is northward. That intensifications in the oval are often associated with polar arcs was an early discovery from the Viking spacecraft imager [Murphree *et al.*, 1987]. It is useful to note here merely that there is a good chance that this isolated intensification, the only one noted during inspection of the Viking images from this orbit, likely correlates with one of the two impulsive southward turnings at 1425 or 1450.

Magnetic activity to accompany this event is not reflected in the magnetic indices nor were definite indications found by inspection of ground magnetograms: the feature appears much smaller than a substorm intensification both in imaging and magnetically. However, it is plausible to infer that the brief southward turning is associated with the release of energy which the UV intensification suggests. The major onset at 1850 may also be associated with sun-aligned arcs in that an intensification occurs on the dayside shortly before substorm onset and likely associated with onset [Lundin *et al.*, 1991].

Figure 7.3 Gain 100 Viking images of Orbit 213, April 1 1986. Transpolar arcs in the evening sector appear to be associated with a brief increase in intensity north of the auroral oval at 1445, seen in both cameras 0 (LBH) and 1 (oxygen). Top left image is before intensification, top right and bottom left during it, and bottom right after it.

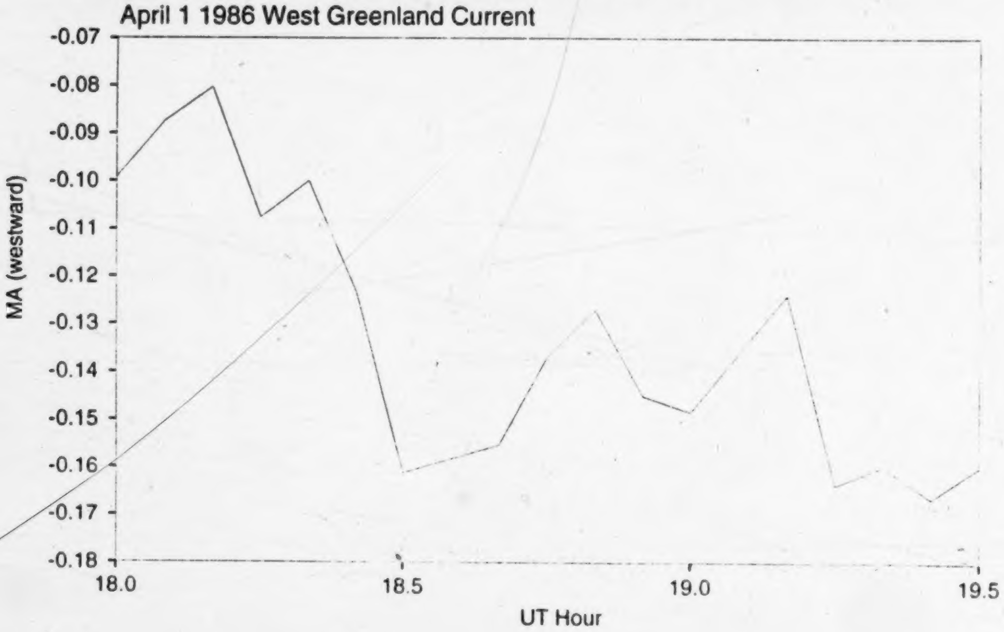


Inspection of initial data plots for global modelling immediately showed two aspects of the current systems before onset: a wide and relatively strong evening sector eastward electrojet is present over Greenland, whereas the morning sector westward electrojet is weak. These two electrojets can be modelled by use of latitude chains and the results of such modelling used to form initial conditions for global modelling.

b. Evening sector electrojet

The eastward electrojet over Greenland was modelled with a simple electrojet using the five-minute quiet-day referenced (using April 4 1986) data supplied by B. Emery. The longitudes of the current system's field-aligned currents were specified as centred on the west coast array: 0° for the upward current and 80° for the downward current, meaning that the prevailing eastward current flow is indicated by the routine as a negative valued current. An approximation routine for inversion of latitude profiles was developed, and using this, the initial parameters for this fit were calculated approximately from the data, based on Z extrema and the amplitude of the X_M variation. The results which follow are thus (apart from the unimportant longitudes) completely automatically generated from the data. Figure 7.4 shows the current over Greenland, which is eastward, as would be expected in the evening sector. The important features here are a doubling of the amplitude of the current about one half hour before substorm onset (i.e. at about 1825 UT) and a relatively constant flow at the new level of about 0.15 MA throughout and after the onset time. It is interesting to note that these changes precede only briefly the precursor activity noted in Viking studies by Murphree *et al.* [1991], which accompanied the disappearance of the transpolar arc. Like the intensifications shown in Figure 7.3, the precursors noted on the Viking (orbit 214) images were at the foot of the transpolar arc.

Figure 7.4 Westward current over Greenland from 18 UT to 1930 UT April 1 1986. This evening sector eastward electrojet has westward current from about -0.08 MA to about -0.16 MA, with current intensification preceding substorm onset.



The electrojet borders, as seen in Figure 7.5, were at relatively high latitude in the evening sector, between roughly 72° and 78° with some evidence of southward motion, particularly after onset, when about 2° of southward motion occurred between onset and 1930 UT. The following series of plots (Figures 7.6 to 7.9) shows the fits and original data involved in the electrojet model. It will be noted that the Y_M component would not be well fit by the simple electrojet model (the model result is zero Y_M and is not shown), but remains at a roughly constant level of about -20 nT at all stations in all plots. The presence of '-D bias' had been noted by Brathwaite and Rostoker [1981] and for the -2 to -4 nT of negative IMF B_z present at 1800 to 1900

UT on April 1, 1986 (see Figure 7.1), their statistical results would lead one to expect to see -10 to -20 nT of bias, which is not inconsistent with what is seen. Whereas in their averaged data there was a correlation between the strength of negative B_z and the Y_M bias, here the bias appears constant while B_z varies. To suggest a physical origin for the Y_M bias, in Figure 7.8 the electrojet model has been modified by the addition of a north-south current in the ionosphere, with its corresponding Region 1/2 currents. The model with the north-south currents added gave almost identical electrojet parameters to that without, with the electrojet from 71.6° to 76.9° and 0.147 MA of eastward current. The north-south current was 0.460 MA towards the north. This result in itself is not particularly meaningful since the length of the system (80°) is somewhat arbitrary, and the ground Y_M perturbation per unit of current decreases with system length (see section 4.d.2.b). Taking an average latitude of 75°, the linear current density comes to 0.20 A/m, comparable to that found in section 6.d in the evening sector, and to that found by Fujii *et al.* [1993], making this explanation for the observed Y_M bias reasonable. This allows a good fit to the Y_M component bias as shown by dashed line in the figure. The geometrical parameters of the borders of this current are identical to those of the electrojet and only the total current in the system has been varied to match the data. Unlike the electrojet perturbation, that arising from the north-south system is not localized in latitude, which is exactly the signature of the observed Y_M bias. Thus it appears that the Y_M bias may be explained as a signature of the Region 1/2 system.

Figure 7.5 Electrojet borders over Greenland from 18 UT to 1930 UT April 1 1986. The evening sector eastward electrojet is at high latitude and shows southward motion.

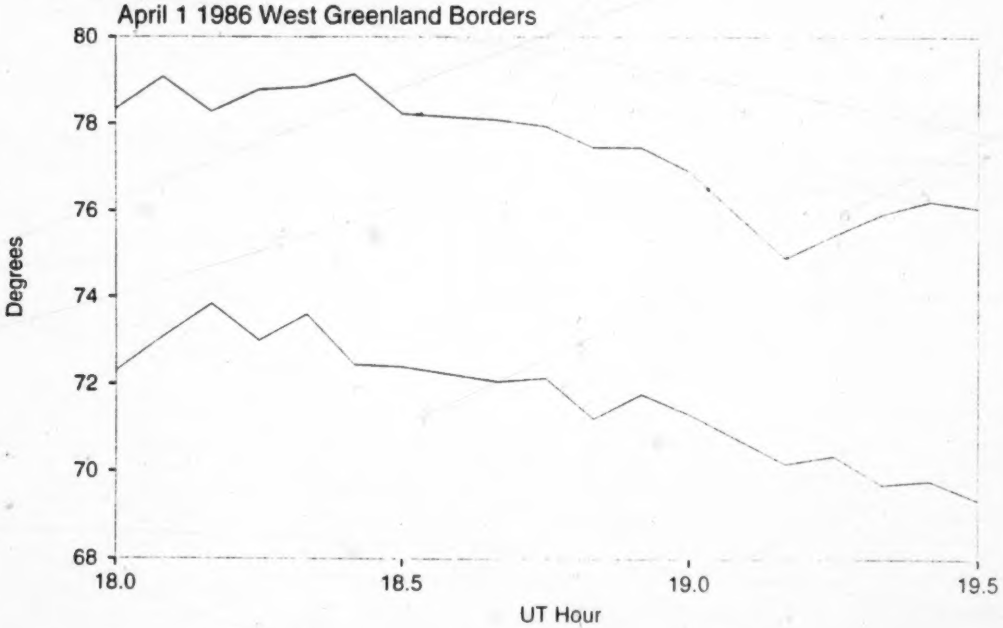


Figure 7.6 Latitude profile for the Greenland array at 1800 UT April 1 1986. Dots are observed X_M component, upward triangles Y_M, and inverted triangles Z. Fits are shown by solid line for X_M and dotted line for Z.

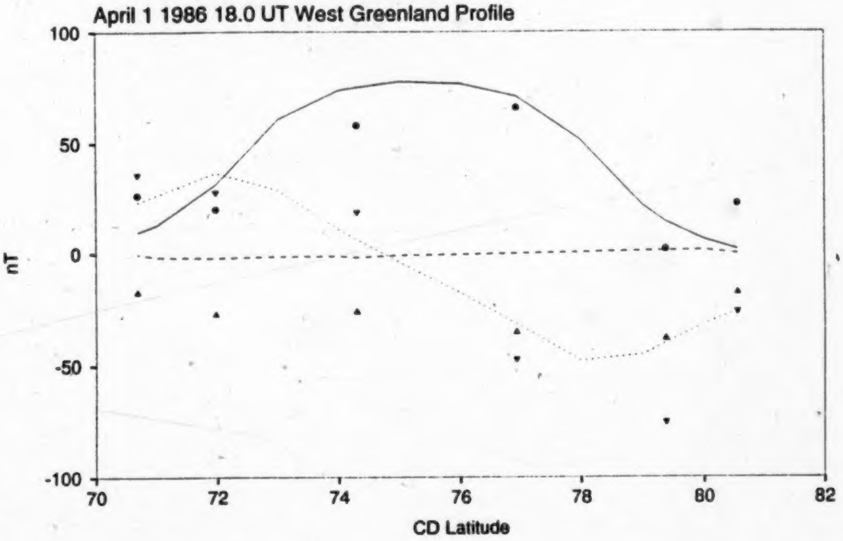


Figure 7.7 Latitude profile for the Greenland array at 1830 UT April 1 1986. Scale has been changed to accommodate larger perturbations.

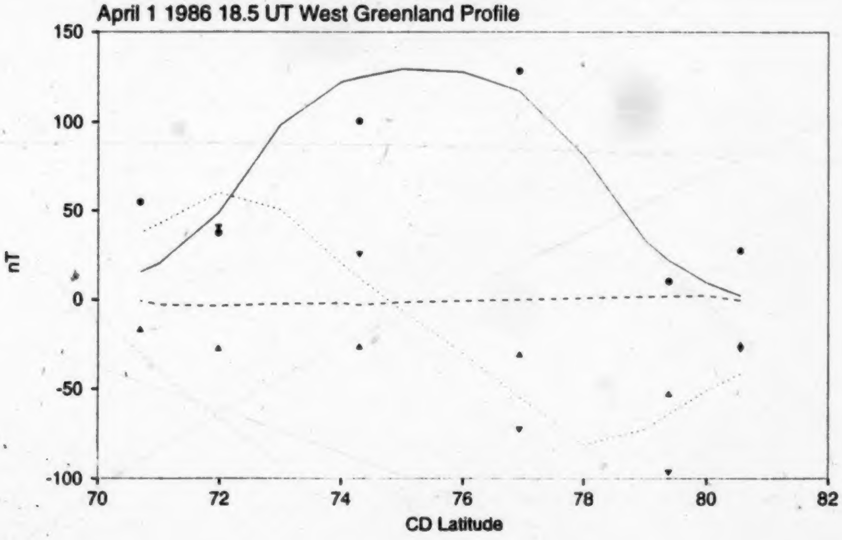


Figure 7.8 Latitude profile for the Greenland array at 1900 UT April 1 1986. Electrojet model is complemented by a north-south system.

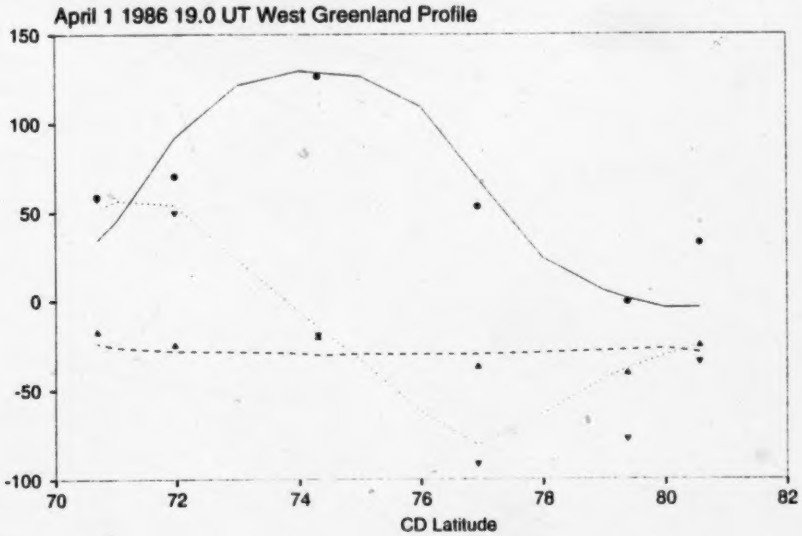
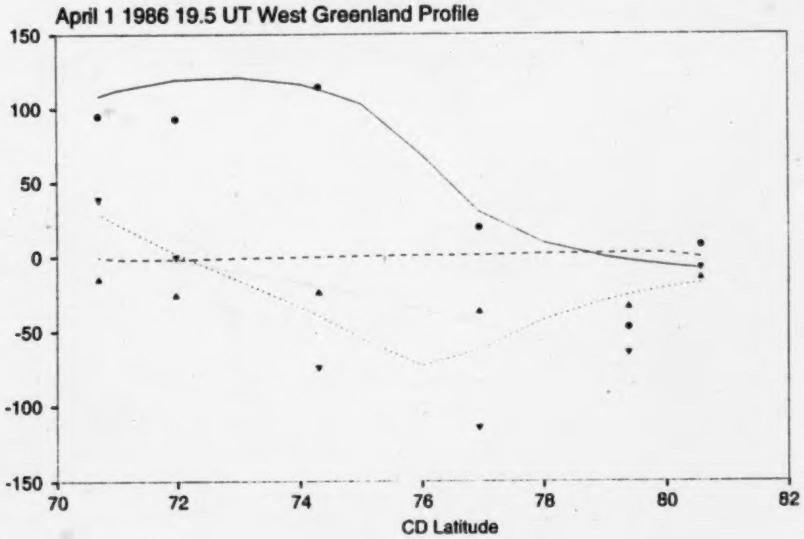


Figure 7.9 Latitude profile for the Greenland array at 1930 UT April 1 1986.



By 1930, as shown in Figure 7.9 southward motion of the electrojet had occurred, with possible widening. Due to the southward motion, the electrojet now extended beyond the southernmost station of the array making that border position less reliable.

Since inferences are being made about Region 1/2 currents, which are easily detected by satellites in low polar orbits, it would be useful to very the conclusions reached here, much as was done in Chapter 6. Unfortunately, in this case the relevant DMSP F7 conjugate passes covering the evening sector near this UT did not reach geomagnetic latitudes below -70° and thus are not suitable for verifying the position or strength of the posited Region 1/2 currents. They do not show any perturbations except possibly at their latitude extrema (data not shown) and this absence supports the far poleward position of the electrojet.

It has been mentioned that a Pedersen system was added to the model to produce the latitude profile of Figure 7.8, with its good match of Y_M component to observation. It is useful at this point to consider the magnetic modelling in the time domain, as illustrated by typical magnetograms from the west coast of Greenland.

Figure 7.10 Observed and modelled perturbations at Frederikshåb, Greenland, on April 1, 1986. Solid lines indicate X_M component: top trace is model. Dashed lines show Y_M component: East-west current model is essentially zero. Dotted lines show Z component: before 18.5 model is top trace, after 18.5 bottom trace. Dots indicate model Y_M component with addition of North-south currents. During this run time step tests were also done and some time steps are missing (notably after 19.0).

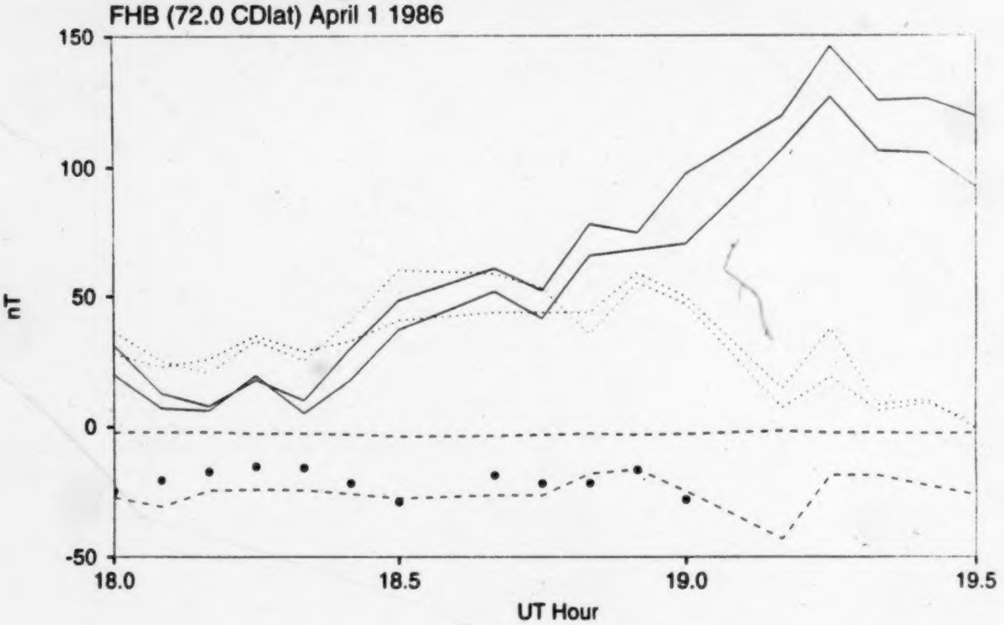


Figure 7.11 Observed and modelled perturbations at Godthåb, Greenland, on April 1, 1986. Traces are as in previous figure. Model X_M is slightly larger than observed, but model Z is smaller in magnitude until 19.0, after which it is above observed trace.

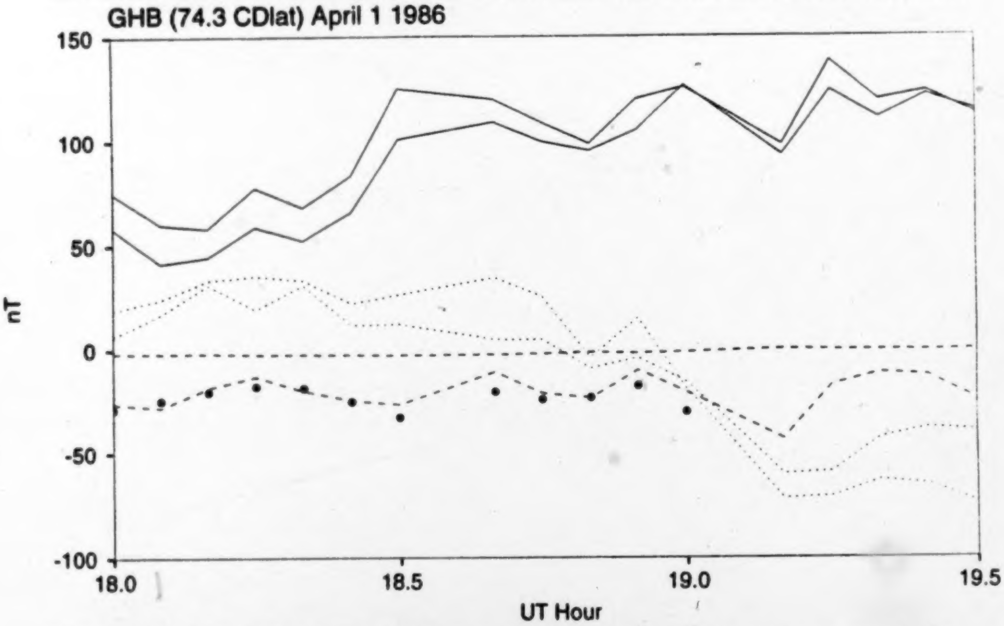


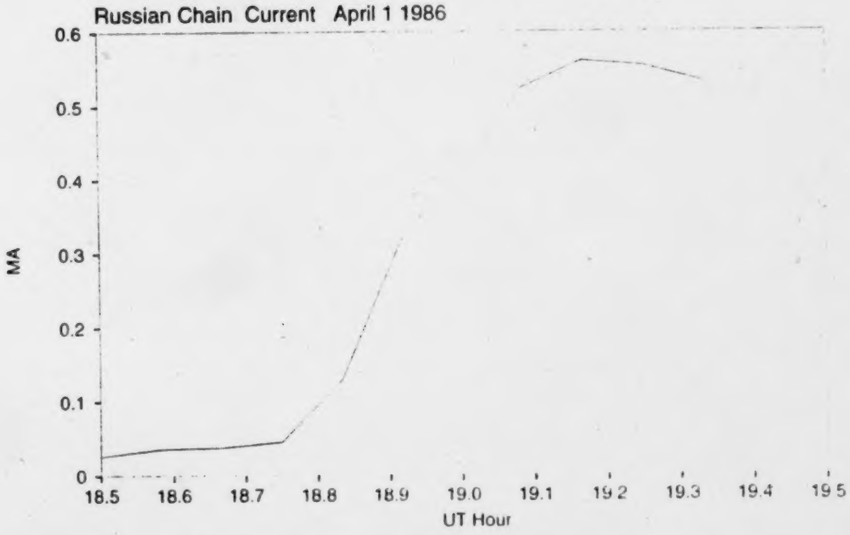
Figure 7.10 shows the perturbations at Frederikshåb (FHB) at 72° centred dipole latitude and Figure 7.11 those at Godthåb (GHB) about 2.3° further north. It may be readily seen that there is excellent agreement between the X_M and Z components at both stations. In the model without Pedersen current, the Y_M component essentially does not vary, remaining at zero throughout the model run. In the model with Pedersen currents added, there is very good tracking of the observed Y_M component by the model at both stations. Before proceeding to discuss use of parameters of the model as a means of efficient representation of magnetic variations, one can compare the general spatial information about the electrojet in the magnetograms and in the border representation of Figure 7.5. At the more northerly station of Godthåb (Figure 7.11) the Z perturbations before 19.0 UT are small and the X_M perturbations larger. This indicates that this station is near the centre of the electrojet, a view supported by noting the positive Z perturbation and smaller X_M at the more southerly station (Figure 7.10). At 19.0 UT, the situation changes and Frederikshåb appears to be near the centre of the electrojet, a view reinforced by seeing negative Z at Godthåb at this time, indicating that an eastward electrojet is now south of it. These deductions are echoed by the modelled electrojet borders shown in Figure 7.5, showing a southward jump of the electrojet at about 19.0 UT. If we now consider the amount of information in a latitude profile (say that of Figure 7.6), we see 6 stations with 3 measurements each, so that there are 18 pieces of information. With a good forward model, including colocated Pedersen currents, the information in such a latitude profile, or in the magnetograms comprising it, may at any time be well represented by only 4 quantities: the northern and southern border latitudes, the east-west current, and the north-south

current[1]. Thus the parametrization in terms of physical parameters is seen to be both accurate and efficient. Much more insight is gained into electrojet behaviour through presentation of these parameters as a function of time, than can be had by inspection of magnetograms.

c. Morning/midnight sector electrojets

The other available chain crossed initially by an electrojet for this event is the Russian chain. An initial model was based on the fit at 1900 UT, and used for fitting automatically at other times. The current over the chain, as computed, is shown in Figure 7.12. The westward electrojet is very weak, with only about 0.03 MA of westward current from 18.5 UT until about 18.8 UT, at which time we are really detecting the enhanced currents of the substorm current wedge. The wedge current rises rapidly to nearly 0.6 MA or about thirty times the initial strength of the westward electrojet (and roughly four times that of the evening sector electrojet).

Figure 7.12 Current over Russian chain from 18.5 to 19.5 UT April 1 1986. Earlier period shows weak westward electrojet, while period after 18.8 UT shows growth and decay of the wedge current.



The electrojet borders may be solved for and the results are shown in Figure 7.13. Originally in the 60° to 62° range, they remain in that range until substorm onset, at which time the poleward border makes a rapid northward leap. From Viking optical data, Murphree *et al.* [1991] found that the borders exhibited similar general behaviour, that is, were at constant latitude before onset, with rapid poleward motion afterwards. In the eccentric dipole system, they found that the borders lay between about 59° and 65°, with a quoted error of only 0.2 degrees. While the equatorward boundary is in agreement with what is found here, the width of about six degrees is much more than that indicated by the magnetic measurements. Murphree *et al.* [1991] point out the existence of brightenings within the auroral oval which they suspect to be discrete arcs within the wider oval which they consider as diffuse aurora. Thus the suggestion may be made that the discrete arcs carry current while the broader diffuse oval does not represent current flow. The measured latitude ranges may further indicate that the Hall currents detected on the ground are located primarily in the southern part of the more diffuse oval, although images do not suggest such a concentration for the discrete arcs. From either of these measurements, the initial latitude range of the auroral oval in the midnight sector is considerably equatorward of that found in the evening sector. At onset time, the Russian chain is very near local midnight, and therefore would be expected to be at a lower latitude, but the difference is larger than canonical values. The westward electrojet is also rather narrow and this may be related to its weakness. To justify the modelling results and to investigate occurrences around onset time, the series of latitude profiles in Figures 7.14 to 7.17 are now examined. At time 18.5 UT the very weak and latitudinally localized westward electrojet is shown through modelling results and observations in Figure 7.14. The X_M and Z components are in excellent agreement while the Y_M component, not represented in the simple model, has irregular variations.

---

[1] Arguably this is essentially the linear current density since we do not specify the east-west extent as a parameter.

Figure 7.13 Borders of electrojet over Russian chain, 18.5 to 19.5 UT April 1 1986. Crossover between initial two points is due to equivalence of north/south border in specifying electrojet and has no physical meaning. Electrojet poleward border moves dramatically northward after onset time of 18.83 (18:50 UT).

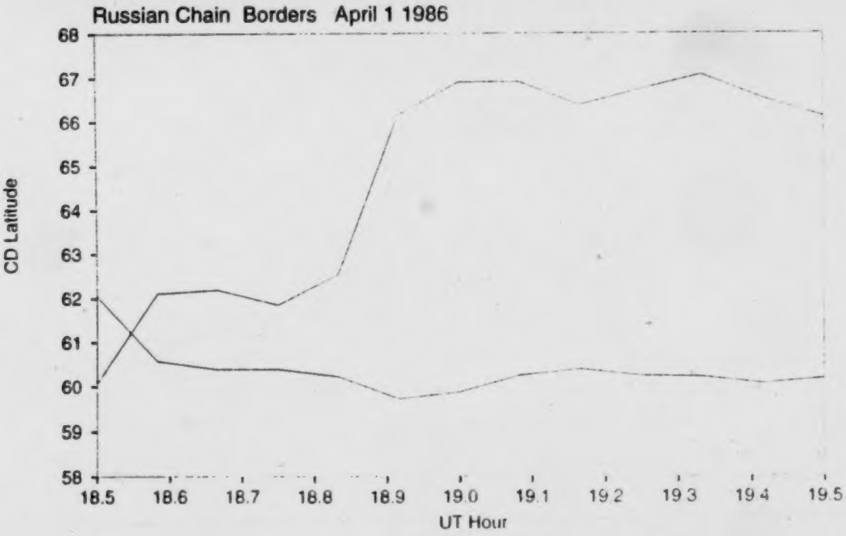


Figure 7.14 Russian chain latitude profile at 18.5 UT April 1 1986. Two stations slightly displaced in longitude at latitude 62° have slightly differing data values.

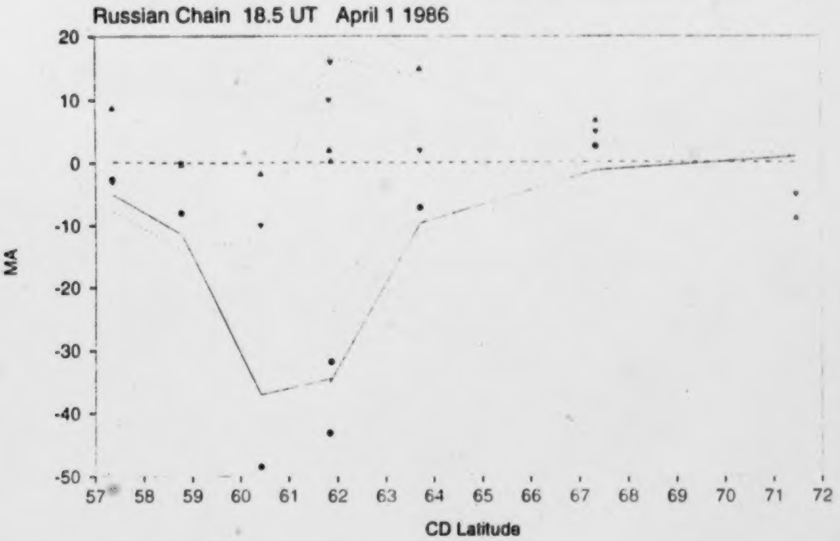


Figure 7.15 Russian chain latitude profile near onset time at 1850 UT April 1 1986.

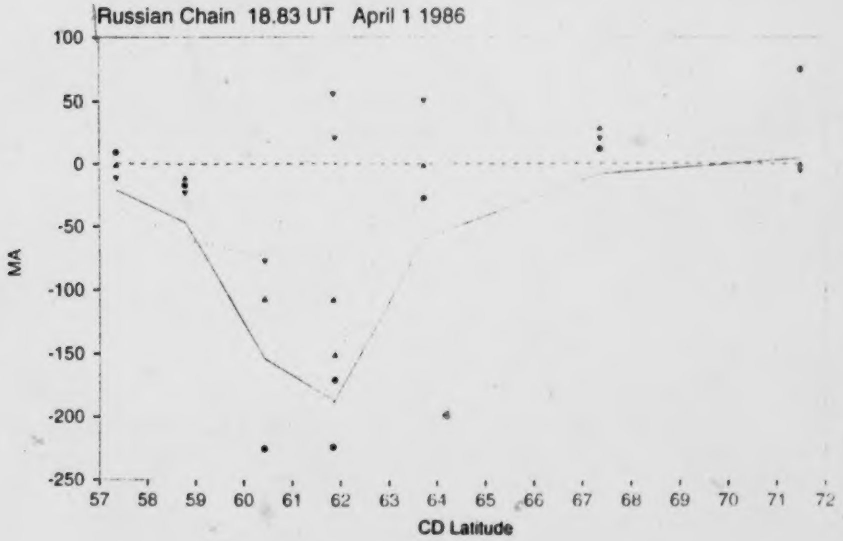
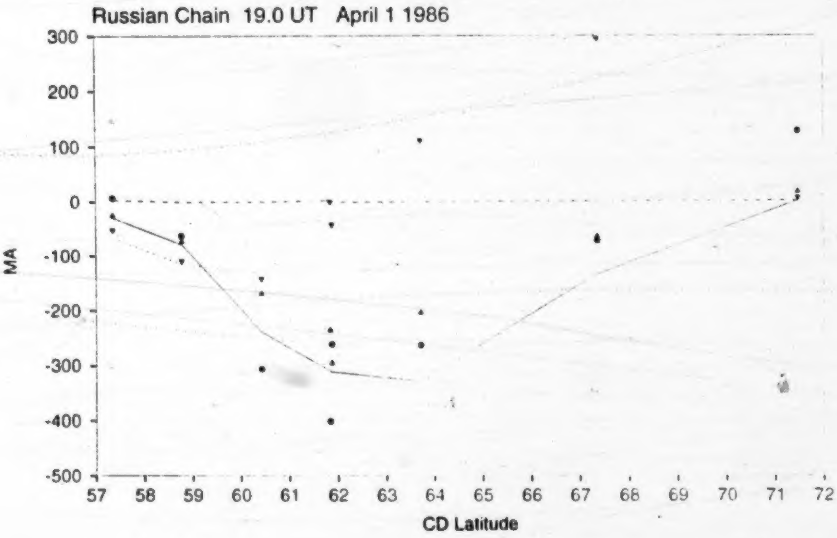


Figure 7.15 shows the profile at onset time of 1850 UT. The westward electrojet, although much strengthened, remains narrow. The large negative Y_M component, marked by upright triangles, follows the X_M component, appearing to indicate electrojet tilt. These negative values of the Y_M component within the electrojet are not modelled by the simple electrojet model (dashed line), which assumes current flow at constant latitude. Only ten minutes later, considerable poleward expansion of the northern border has occurred as may be seen in Figure 7.16. The continued tilt of the electrojet is still indicated by the Y_M component following the X_M component, which appears to be the case at high latitudes (see points at 67°) also. Figure 7.16 illustrates that the poleward border is still at a northerly position as indicated particularly by the Z fit and Z maximum, although the strength of the X_M component at 67° seems diminished. Yet later the basic situation has not changed much although there are deviations in individual components, particularly at the northern border (Figure 7.17). While the more southerly stations have a very good fit to perturbations, the northernmost, Heiss Island, shows increasingly positive perturbation, not well modelled. Inspection of original one-minute data supplied from the Kara Sea chain of the Russian Arctic and Antarctic Institute [A. Kotikov, personal communication, 1996] shows that there is indeed significant perturbation at this station at these times. While Heiss Island is known to have anomalies, the most likely explanation for the failure to model is that the actual physical situation is more complex than that of the simple single-component electrojet model used here.

Figure 7.16 - Russian chain latitude profile at 1900 UT April 1, 1986. This profile, ten minutes after onset, shows the significant poleward expansion of the northern border.



The most likely situation is one of bifurcation of the electrojet, and indeed the more northerly portion appears to carry predominantly eastward current as indicated by the positive X perturbation at Heiss Island. The inference that there are two regions of current flow across the Russian chain is borne out by the study of conductivity above this chain by Kotikov and coworkers [1995]. On 5 dates, including earlier (1654 to 1701 UT) on April 1 1986, precipitating particles observed by DMSP were used in conjunction with an atmospheric model to compute the Hall conductivity. It was found that in addition to the enhanced conductivity in the 'traditional' auroral zone between latitudes 64° and 68°, there was a poleward region of enhanced conductivity in a narrow band found between latitudes 70° and 73° (corrected geomagnetic). While it was generally inferred that this poleward band carried a poleward filament of westward electrojet, the April 1 event was not discussed in detail as to magnetic signatures in the article. Kotikov [1996, private communication] agrees that there is evidence for eastward electrojet, and indicates that there is nothing wrong with the data. The five-minute data, as modelled by the AMIE technique (see below) also indicate a weak eastward electrojet at this location.

Figure 7.17 Russian chain latitude profile at 1930 UT April 1, 1986. Poleward border is in roughly the same position as at 1900 UT and total current has decreased, resulting in smaller perturbations.

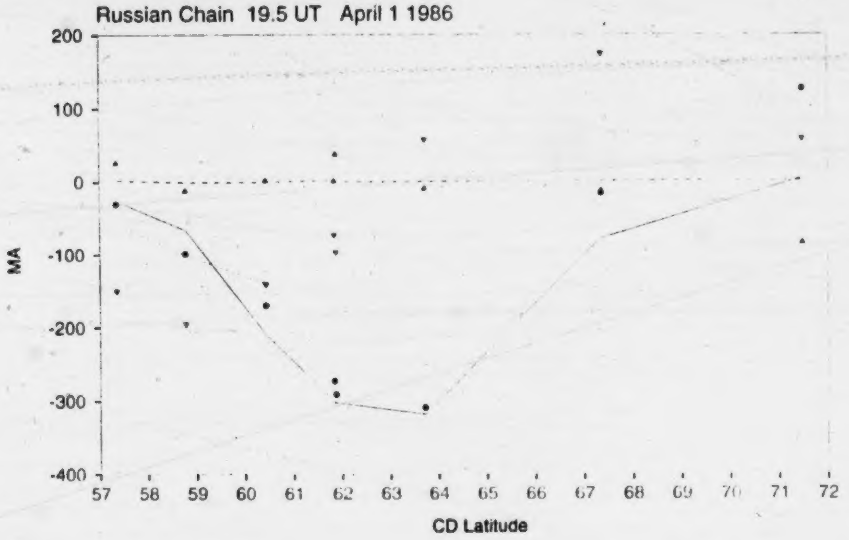
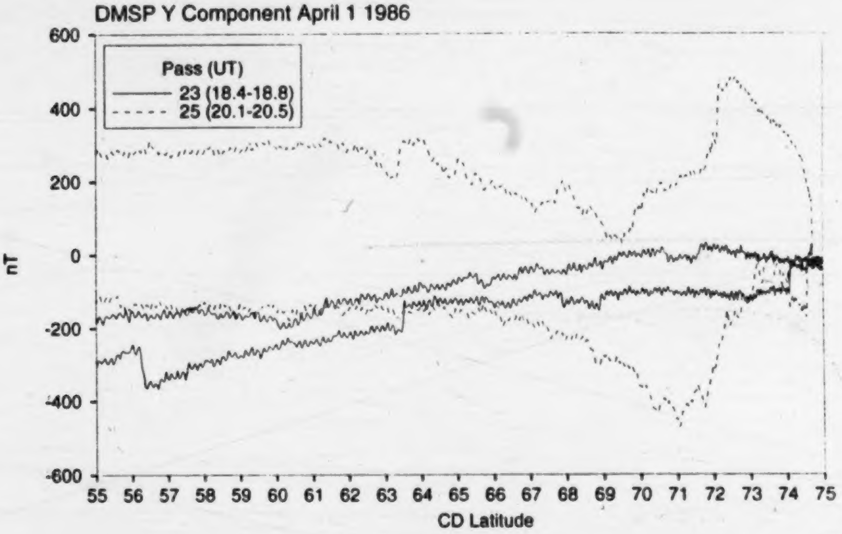


Figure 7.18 DMSP F7 Y component on April 1 1986. Pass 23 (solid curve) shows no significant electrojet perturbations. Pass 25 (dashed) shows westward perturbations in midnight sector (top) from about 62° to 72° and in morning sector (bottom) from about 65° to 73°.



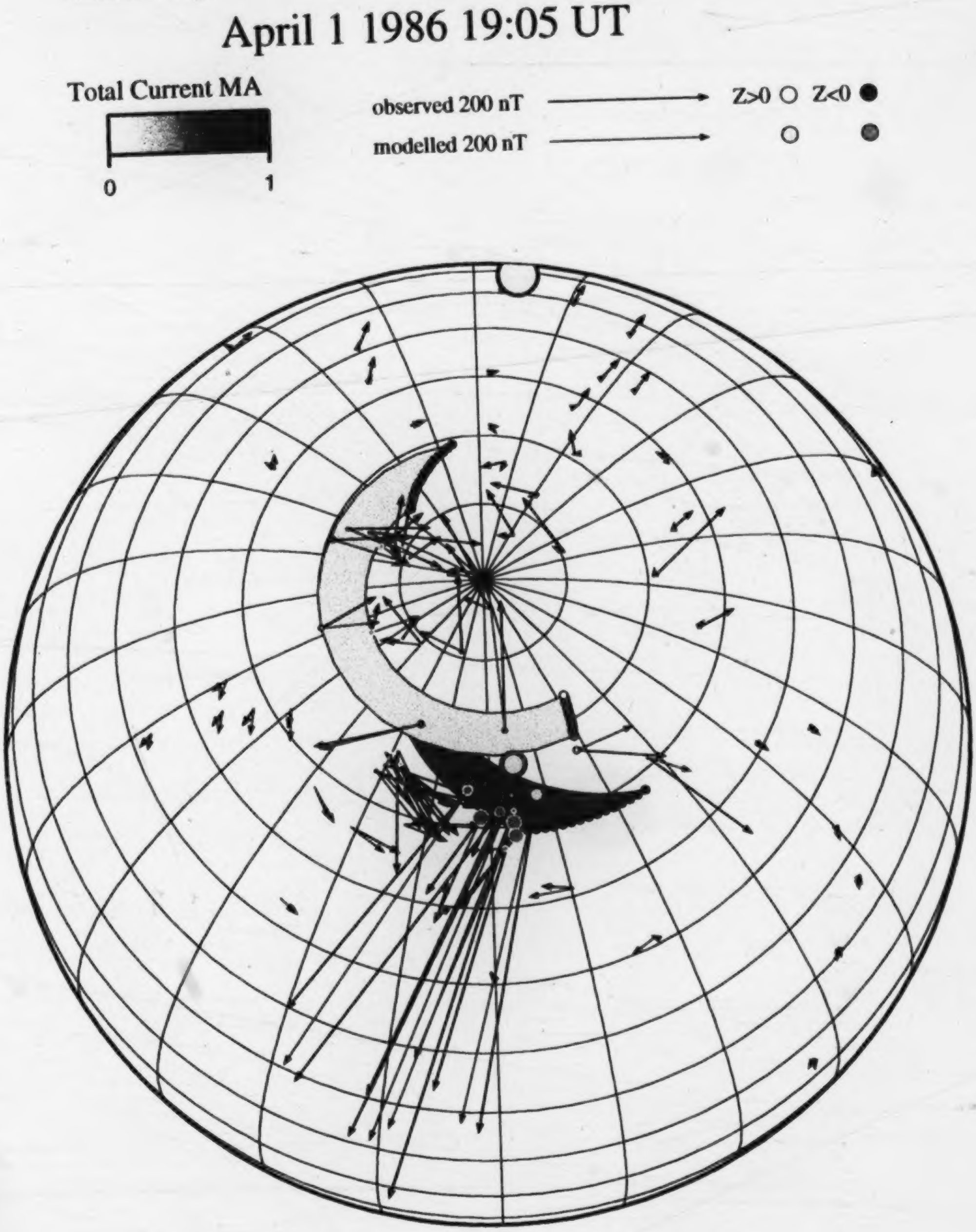
The increase in strength of the westward electrojet during this event is notable. The morning/midnight sector Region 1/2 systems associated with the westward electrojet also change remarkably during the time interval containing the onset. Figure 7.18 illustrates this through the DMSP F7 Y component on a pass from 18.4 to 18.8 UT, just before onset, and one after onset at 20.1 to 20.5 UT. The first pass was during the period modelled and it has been pointed out that the westward electrojet was extremely weak. The lack of perturbations seen on this pass supports this conclusion in that the Region 1/2 currents are either so weak as to be unnoticeable on the plot, or may have been located north of the northernmost point (ca. 75°) on the spacecraft trajectory. The latter explanation is unlikely since the electrojet borders before onset are at roughly 62° and the electrojet thus should have been traversed by the spacecraft. Although the second pass is after the modelling period, inspection of the AE index in Figure 7.2 shows that activity continued at an elevated level during the pass, possibly with at least one more onset having occurred. That strong and nearly identical signatures of north to south Pedersen flow in the morning and midnight sectors have appeared is obvious. Similar dramatic changes in DMSP perturbations were noted by Rich *et al.* [1990] and associated with variations in IMF direction during a major geomagnetic storm in February 1986. In that case both morning and evening sector perturbations varied significantly, and were observable since the auroral oval was very extended and the spacecraft passed through the auroral oval in both local time sectors. Here, as noted above in discussing the addition of a Pedersen system over Greenland, DMSP signatures from the evening sector are not available at suitable latitudes. Given the small change in the already strong evening sector eastward current flow inferred above, it seems unlikely in this case that the evening sector Region 1/2 currents show the same degree of dramatic increase as that noted in the morning sector. One final point is that the enhanced currents seen during the second pass appear to be unbalanced in the midnight sector, with net upward current. As before, the linear current density may be calculated from $i_z = \frac{\Delta B_y}{\mu_0} = 8 \times 10^{-4} \Delta B_y$ in A/m if B_y is given in nT, and the total current in the Pedersen system is the product of this linear current density by the east-west extent of the system. There appears to be about 400 nT of level shift, corresponding to about .3 A/m linear current density. What relationship, if any, this current bears to the anomalous perturbations at Heiss Island, cannot be determined from these data.

d. Three-dimensional modelling

In proceeding to three-dimensional analysis, it may be noted that the current flowing in the Greenland and Russian chains is quite well established through latitude profiles. The borders of the electrojets are also essentially known, although there is some uncertainty about the northern border of the westward electrojet. It is clearly established that the evening sector eastward electrojet is strong and that the westward electrojet is weak at pre-onset times, to the point that the perturbations of the substorm current wedge appear from essentially nothing. This being the case, an initial attempt to model using only an eastward and westward electrojet was undertaken.

Representative results are shown in Figure 7.19. It can be seen that the evening sector is well represented by an electrojet at roughly 75° latitude, but that it crosses the midnight sector to respond to the large positive perturbation at Heiss Island. The westward electrojet, here as essentially an isolated substorm current wedge, angles upward in the late evening sector to the early morning sector and the upward current at its western end is within the region of the eastward electrojet. As it has a large current flowing upward there, it provides perturbations south of it which reflect well the primarily eastward perturbations observed in Scandinavia, as well as in a general sense those in western Europe. On the other hand, perturbations in the mid and early morning sectors leave no doubt that there are in fact local effects from a global westward electrojet, and perturbations in the polar cap are not well reproduced. Having the eastward electrojet traverse the region of upward current associated with the substorm current wedge is not physically likely (although not impossible). Further, the model does not represent morning sector features well due to the fact that the only westward electrojet is constrained observationally to lie in the midnight sector to represent the strong current wedge. Although a number of features are correct in this two-electrojet model, a better model was desired. The forward model was modified to have a westward electrojet separate from the substorm current wedge, to incorporate Pedersen currents likely responsible for polar cap perturbations, and to not allow penetration of the eastward electrojet past the substorm current wedge. This latter modification was expected to be at the expense of reproducing perturbations at Heiss Island, but this station is felt to be anomalous in any case.

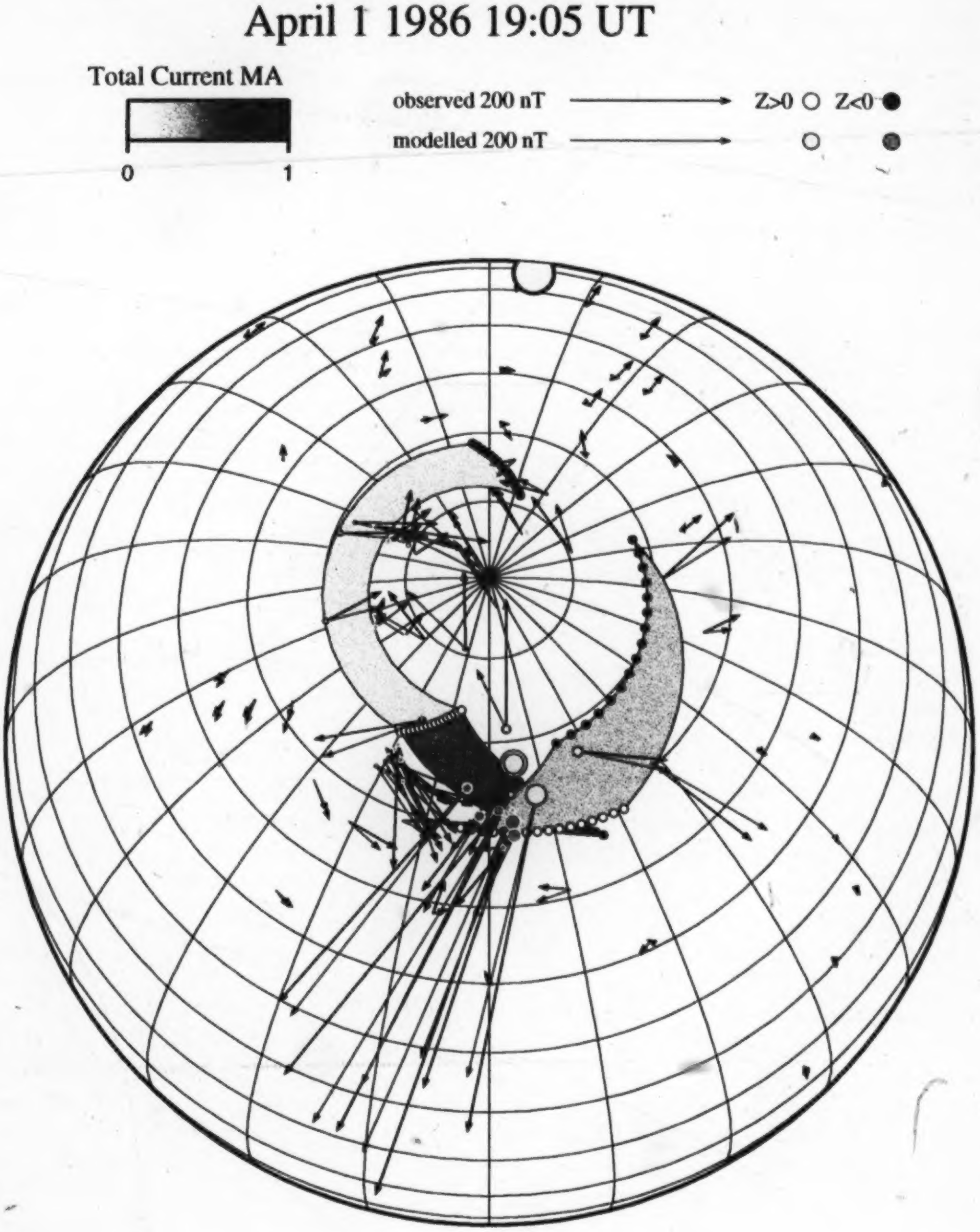
Figure 7.19 Global modelling results using eastward electrojet and current wedge.


April 1 1986 19:05 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●

The result of these modifications is shown in Figure 7.20. Morning sector perturbations are well represented, with downward current flow from roughly 0900 MLT to 0100 MLT feeding a widening electrojet. Midnight sector perturbations are well modelled by a wedge which is tilted northward (that is, from southeast to northwest while following the direction of Hall current flow) at a relatively steep angle, producing electrojet perturbations in the X_M and Z coordinates near it. Consistent with this are the Y_m signatures normally associated with field-aligned current, found south of its associated upward current flow. The tilt of the electrojet is based on both X_M and Z perturbations, with a change in polarity of Z visible across the electrojet at various points. Thus the correspondence of X_M and Y_m perturbations, noted above in discussing results from latitude profiles, is indeed due to electrojet tilt and not to Region 1-Region 2 effects. The direction of tilt is the opposite of that found from AMIE modelling by Lu *et al.* [1997]. As noted in Chapter 4, the latitudinal resolution of techniques based on spherical harmonic decomposition is not high. Thus the results presented here are felt to better represent the true configuration of electrojets than are those given by Lu *et al.* [1997].

The evening sector electrojet allows perturbations in Greenland to be well modelled, but has been constrained to meet the wedge electrojet at its point of upward current flow, thus as expected not allowing the positive X_M perturbation at Heiss Island to be modelled. Due to the colocated Region 1/2 field-aligned current flow, polar cap perturbations are relatively well modelled. The major objection to this solution is that there is a region of overlapping upward and downward current flow in the vicinity of the juncture of the westward electrojet and substorm current wedge. While this may be understood in that the wedge current is stronger, so that there is in fact net downward current flow in this region, it might be desirable to constrain the two systems to have one junction, much as was done in the previous chapter.

Figure 7.20 Global modelling results using two electrojets and a current wedge, including associated north-south currents


April 1 1986 19:05 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●

e. Summary

This event has been seen to conform rather well to the 'canonical' pattern of an isolated substorm. Southward turning of the IMF is followed by a growth phase and a substorm expansive phase. The growth phase in this case is mainly manifest in the strengthening of the evening sector eastward electrojet, which is found at high latitude. The expansive phase accompanies northward turning of the IMF and is marked by a current wedge near the midnight sector.

The following phenomenological details have been noted:

1) strong evening sector eastward electrojet at high latitude

2) weak morning sector westward electrojet at lower latitude, becoming stronger after onset

3) latitudinal motions of the electrojets at onset time, including poleward expansion of current wedge

4) turning on of morning sector region 1/2 currents after onset.

5) significant northward tilt of the current wedge electrojet.

The following conclusions are related to modelling issues:

6) region of juncture of westward electrojet and wedge is complex

7) inclusion of region 1/2 currents allows representation of Y_M bias near evening electrojet, and of polar cap perturbations.

Chapter 8. CDAW 9C - May 3, 1986

This period has been relatively well studied and was selected for inclusion here due to data availability since it had been included in the CDAW 9 analysis suite. The entire CDAW period is from 00 UT to 12 UT and features several substorm intensifications, but magnetospheric activity was already at a high level well before this time. The earlier part of the CDAW period, featuring a major onset at 0111 UT, has been studied by Pulkkinen *et al.* [1992] and by Baker *et al.* [1993] with emphasis on satellite data and an interpretation in terms of near-Earth neutral line theory. These authors pointed out a discrepancy in times of onset inferred using satellites and the AE index. Features in the morning sector were examined by Connors and Rostoker [1993]. The latter part (0900-1030 UT May 3) features an onset examined by Hones *et al.* [1987] in the context of plasmoid models and using ISEE data, and revisited by Mishin *et al.* [1997] using the MIT method concentrating on ground magnetic observations. The overall period is one of magnetic storm, commencing about 07 UT on May 2, with maximum negative values of Dst (see Figure 8.13) of about -80 nT shortly before 0 UT and at 07 UT on May 3 (near the time of the Hones study), and recovering into May 4. The ring current ion population during this storm [Stüdemann *et al.*, 1987] was studied using Viking and AMPTE/CCE particle data. Discussion here will be rather brief and will focus on the use of indices and on regional modelling to answer specific questions about activity in the time period 0 to 3 UT on May 3. Some modelling was also done for May 2 to establish the conditions leading to the main activity discussed.

Figure 8.1 Auroral electrojet indices from 18 UT May 2 to 12 UT May 3 1986. Upper trace is AE, middle trace is AU, and bottom trace is AL. Period considered here is from about 0100 to 0300.

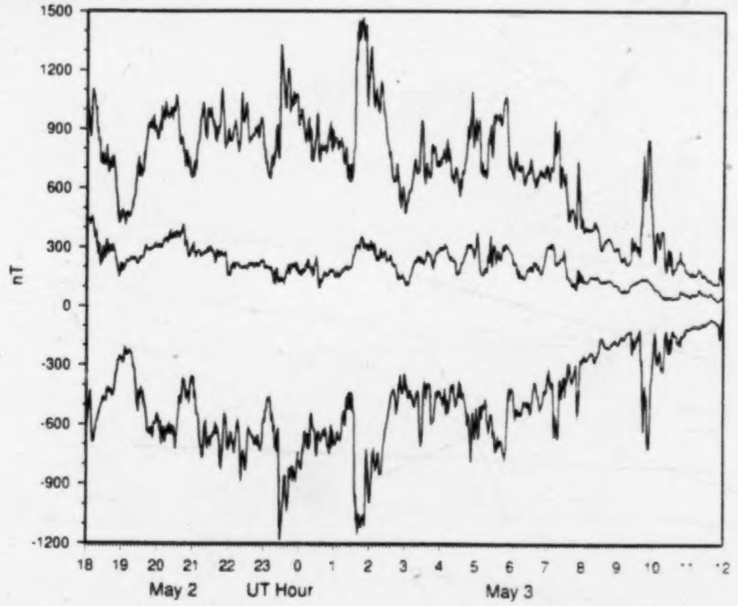
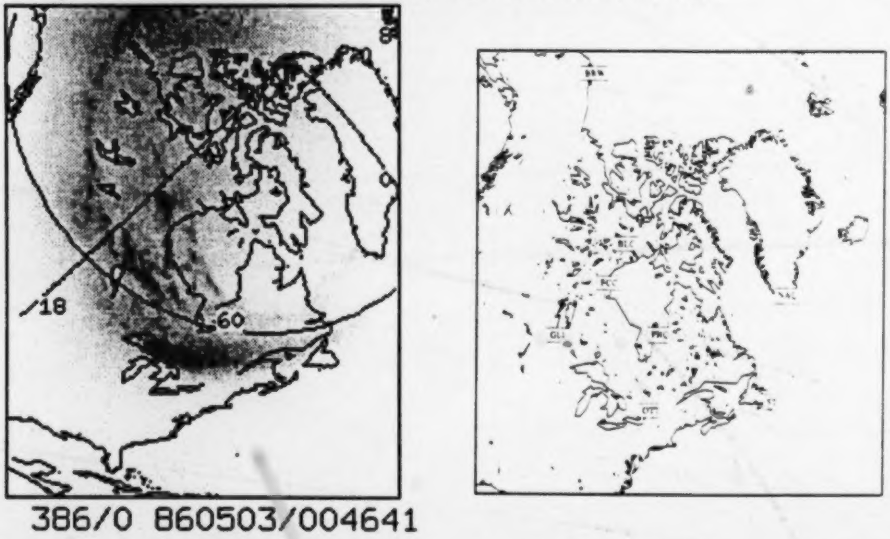


Figure 8.1 gives an overview of conditions on May 2 and 3 1986 through the electrojet indices[1]. The entire sequence of activity leading to the high activity level shown started with a sudden commencement at about 2010 UT on May 1, 1986 [Baker *et al.*, 1993]. Solar wind data are not available for the entire storm period and most particularly are not available during the actual CDAW 9C period due to IMP-8 being in the magnetotail. For this reason the solar wind context will not be discussed here. Several sharp onsets are recorded in AL for May 2 (partially shown in the figure; note onset at 23.2 UT), with AL levels of -1000 being common in the 1400-1800 period on that day. 1800 May 2 to 0130 May 3 is typified by AL of about -600, with the noted sharp onset to -1200 at 2325 May 2. The onset indicated by AL at 0130 May 3 is the first shown by the index during the CDAW period. This specific enhancement of activity ends by 0300 with a return to AL of about -500, still indicative of very active electrojets. Other onsets occurred later and one beginning at 0919 was studied in detail by Hones *et al.* [1987] as mentioned above. After about 1030, relatively quiet conditions had returned, although AL was then still about -100 compared to lesser values (about -50) at this time on the nearest quiet day (May 1). Nevertheless, given the high activity levels during the period of interest, the AE behaviour suggests that one would not go too far astray in using observed magnetic field values from after 1030 UT for baselining. In practice, one-hour smoothed values for 11 UT were used for many of the stations used in this study. The perturbation data were the result of subtracting these values from those of the respective components in the original data files. Where possible, the perturbation data were checked against the May 1 quiet day, or that day was smoothed and subtracted off to give perturbations.

Figure 8.2 Conditions prior to major onset on May 3 1986. Camera 0 image from Viking. Inset shows some station locations as discussed in the text.



---

[1] ftp://ftp.ngdc.noaa.gov/STP/GEOMAGNETIC_DATA/INDICES/AURORAL_ELECTROJET/ONE_MINUTE/ae86.1-6 is the source for AE for May 2; CDAW CD-ROM for AE, AU, AL on May 3

a. Background

The cited study by Baker and co-workers found that there was a discrepancy in the positions of auroral brightening and maximal electrojet current in the onset at 0111 UT. The brightening occurs at roughly 2200 MLT in the Ottawa-Poste de la Baleine (PDB) region and at latitudes close to Ottawa, while the current appears from the magnetograms to maximize near (although south of) the more northerly station. Using magnetic latitudes of 56° and 66° for these stations, it is claimed that the electrojet is 10° north of the brightenings as imaged by Viking. As seen by the fact that the onset indicated by AL is at 0130 and not 0111 (although AE does decrease slightly at 0111), the index must again be taken as not accurately showing expansive phase activity. This situation is comparable to that found for the onset seen at Gillam after 0600 during the CDAW 9B event (section 6.j).

Given the high activity level prior to the start of the CDAW interval at 0000 May 3, it is not surprising that imaging reveals the auroral oval to be wide and active as shown by Figure 8.2. The early evening sector auroral oval is also seen to be steeply rising northward, while in the morning sector it nearly follows lines of (EDFL) geomagnetic latitude. Corresponding DE-1 data and rectified Viking data are presented in pseudocolour by [Baker *et al.*, 1993], but it is not clear from the DE-1 images that the oval does rise in latitude as one approaches the dayside. The effect is, however, present in the rectified Viking images shown there. Auroral activity is also visible at low latitudes in the late evening sector, well below 60°. This activity appears to be generally diffuse, but near the northern border of the oval, discreet and perhaps periodic forms are seen. Discrete forms are also seen within the oval.

To better illustrate the general situation, and to show the onset at 0111, further Viking images are shown in Figure 8.3. The evening sector auroral oval is doubled and very wide, at least 20° in north-south extent, and the morning sector oval is also doubled and wide. The overall oval position is still far to the south, approaching the latitude of Ottawa (about 45° geographic, 55° magnetic) in the evening sector. This is consistent with the position of diffuse auroras as noted in Figure 8.2. The foreshortening in this region of the globe becomes worse as the image sequence progresses. Thus, having established the far southward diffuse oval position through the earlier images is important. One image (at 011029) immediately precedes onset, and in this image the oval is seen to be clearly doubled, with two bands of luminosity visible south of Greenland, in the vicinity where the onset brightening would soon take place. The final image shows the situation about 6 minutes after onset, with marked brightening due most likely to heavy electron precipitation. Besides the worsening foreshortening, the images also suffer from a changing field of view which no longer includes the early evening sector nor the most equatorward part of the auroral oval in the late evening sector.

Figure 8.3 Vignetting-corrected Camera 0 Viking images on May 3 1986. Image at 005625 (upper left) shows wide auroral oval in evening sector. This and discrete morning auroral forms are seen at 010347 (upper right), while evening sector is not easily visible in immediate pre-onset image at 011029 (lower left). Note doubling of auroral oval in eastern Canada. Lower left image shows brightening associated with 0111 onset.

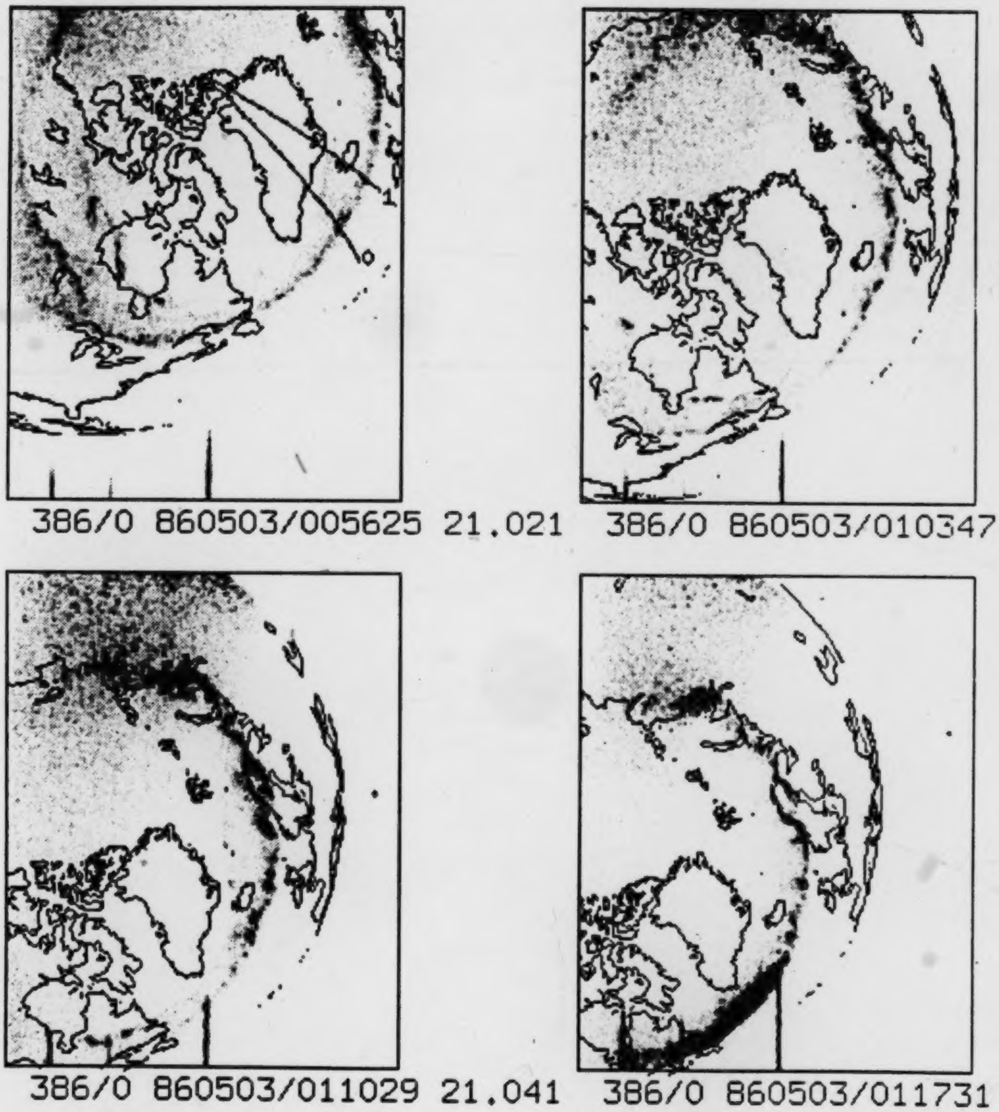

1
0
386/0 860503/005625 21.021
386/0 860503/010347
386/0 860503/011029 21.041
386/0 860503/011731

Figure 8.3 also shows discrete features at the poleward border of the oval and these merit some discussion. It should be noted that the image taken at 005625 shows a band of luminosity paralleling the east coast of Greenland between Greenland and Iceland. The band is angled poleward from the discrete oval and can be as much as about ten degrees poleward of it, and is located from roughly 2330 MLT to 0130 MLT This feature is no longer obvious at 010347 but has brightened, having drifted somewhat eastward (and perhaps brightened at its eastward end which is now located at roughly 0400 MLT), just before onset in the 011029 frame. This brightening, and the fact that there are discreet and periodic forms at this position, were noted by Connors and Rostoker [1993]. As they noted, these forms faded after onset as inspection of the 011731 frame will show. Also of significance is the presence of discrete features at the poleward border in the late evening sector in the pre-onset frames. These are located near the eastern shore of Hudson Bay almost directly above the magnetic station at Poste de la Baleine[2] (see geographical locator inset on Figure 8.2). The presence of these features is associated with features on the station's magnetogram, which is shown in Figure 8.4. Starting shortly before 0100, there is a decline in the X component, and shortly after a further, short-lived decline which has a Z perturbation associated with it. Before the onset time (here marked at 0112), there is a systematic decline of nearly 100 nT in the X component. This decline is clearly associated with the luminous features at the poleward border of the auroral oval, and the variable character with either their motion or other type of fluctuation. The relatively large magnitude of the ensuing, onset-associated decline in X and corresponding rise in Z suggest a westward current immediately south of the station as inferred by Baker *et al.* [1993]. The electrojet strengthens at about 0130 and seems to be further south as indicated by the increased Z component. Y is significant during this period, comparable in magnitude to the X perturbation although positive in sign.

In hopes of clarifying the situation at the poleward border, recourse was had to conjugate data. The station Siple is very nearly conjugate to Poste de la Baleine in the centred dipole system, having coordinates of (352.5,-65.24) to the latter's (350.4,66.1). Perturbations at Siple in the local magnetic system are shown in Figure 8.5. There is general agreement between the behaviour of the corresponding components of the fields at these two stations once the reversal of the eastward and Z components is taken into account. However, there seems to be more evidence of evening sector eastward electrojet prior to onset, the 0112 onset (perhaps starting at 0111 at this station) is of slightly lesser amplitude in the northward component and considerably so in the Z component. There is a definite second strengthening of the electrojet at about 0130 as at PDB, but of lesser amplitude and without much change in the Z component. The eastward component shows a curious reversal during this intensification, attaining over 800 nT shortly before 0200. Later intensifications appear which are not observed at PDB. It is difficult to interpret the lessened Z perturbations as evidence that Siple is closer to the intensified electrojet given the accompanying decrease in the amplitude of the northward component.

---

[2] also known as Great Whale River and now officially as Kuujjuarapik: symbol GWR. PDB. PBQ

Figure 8.4 XYZ magnetic perturbations at Poste de la Baleine on May 3 1986. As in previous chapters, X is indicated by a solid curve, Y by a dashed curve, and Z by dots. Onset time is indicated by a vertical line at 0112 UT.

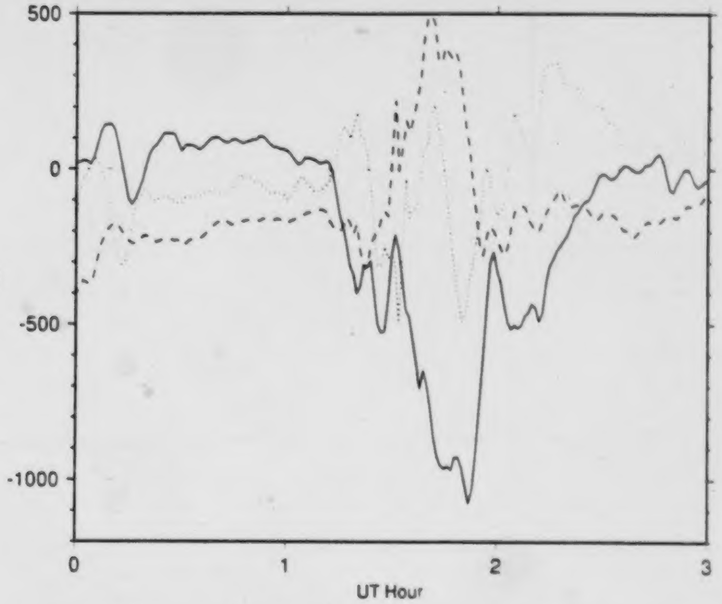
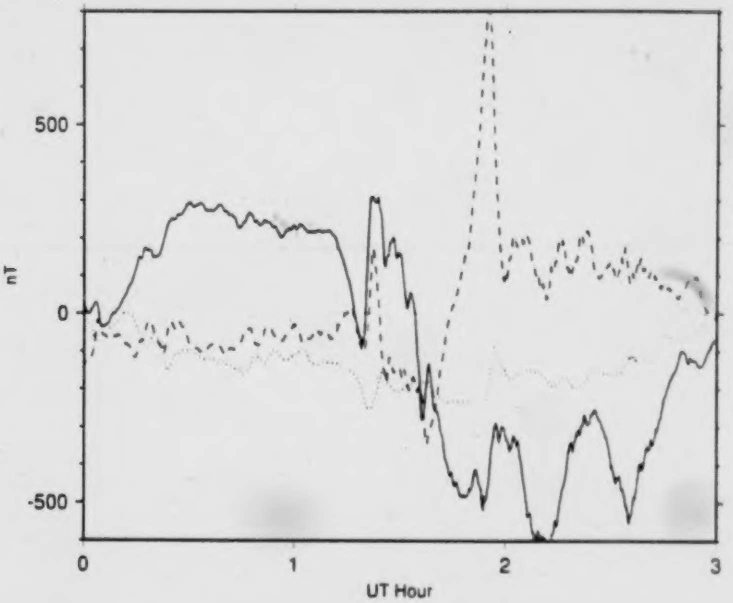


Figure 8.5 Perturbations in local magnetic system at Siple, Antarctica, on May 3 1986. Onset time of 0112 is shown by a vertical bar.



To further explore conjugacy for this event, data from the known close conjugate pair Syowa-Husafell may be used (recall discussion in Chapter 3, and in particular Figure 3.9). The local magnetic northward and Z perturbations are presented (using pulsation data with 10 second sampling period) in Figure 8.6. The local eastward components are not shown but show excellent anticorrelation as would be expected for this component. The northward perturbations show good agreement although clearly not as good as that which obtained during the CDAW 9B event discussed in Chapter 6. The Z traces are somewhat puzzling in the similarity of the significant Z values. The fact that there is a Z value normally indicates that the electrojet is not centred on the station. In the case of Husafell this correctly (see Viking images) indicates that the electrojet is south of the station by being positive. However, in the conjugate case it is also indicated that the electrojet is south of the station and this is not what would be expected from good conjugacy. The two Z traces continue to be similar throughout the CDAW 9C period (data not shown) including a large change of polarity at about 6 UT. It appears that the electrojet is offset in latitude in the northern and southern hemispheres, being (before 6 UT) equatorward of Husafell yet poleward of its conjugate. If the latter interpretation is correct, then this would imply that the current being sensed is very narrow. This conclusion is not inconsistent with one which could explain the differences between the PDB and Siple results despite the relatively small latitudinal separation of their conjugate points. Another useful fact to note from Figure 8.6 is that the 'onset' time as recorded at both Syowa and Husafell is close to 0142, and that there is a marked positive Z perturbation at Husafell at this time. Similar events are seen near this time at the nearby stations Isafjordur and Tjornes. An interpretation in terms of narrow current at the northern border of the electrojet moving northward until over the station and then stopping seems tenable: this will be examined in more detail below. Comparison of the perturbation at Icelandic stations shows the general features expected of stations near the northern border of the electrojet and the data seem suitable for use in modelling.

While the general interpretation of Baker *et al.* [1993] that prior to 0111 a growth phase was underway is undoubtedly correct, there is evidence of a small onset at roughly 0030. This is before Viking imaging commenced, and the earliest Viking images do not show Europe, which is where some onset signatures are evident. The onset appears to have taken place within the active morning sector electrojet at about 0200 MLT. There is some evidence for a minor injection from the satellite 1984-037 which is at a local time of about 0300, and in geosynchronous orbit at about -8.6 degrees in centred dipole coordinates (data not shown). Stronger growth phase signatures at satellite 1986-019, which is very close to midnight, are seen in Figure 8.7. At this location there appears, at 0030, to be an enhancement in the lower energy channels while the higher energies drop out. Ground stations west of midnight (i.e. in North America and Halley Bay) do not show signatures of the putative onset, but it is marked by a significant enhancement of the electrojet as observed in Scandinavia. Figure 8.9 shows perturbations at Muonio, and it is seen that a decline in the X_M value started at about 0030.

Figure 8.6 Local magnetic northward and Z components at Syowa and Husafell. Heavy traces correspond to Syowa and lower trace pair is the northward perturbation. Vertical bar marks 0142 UT.

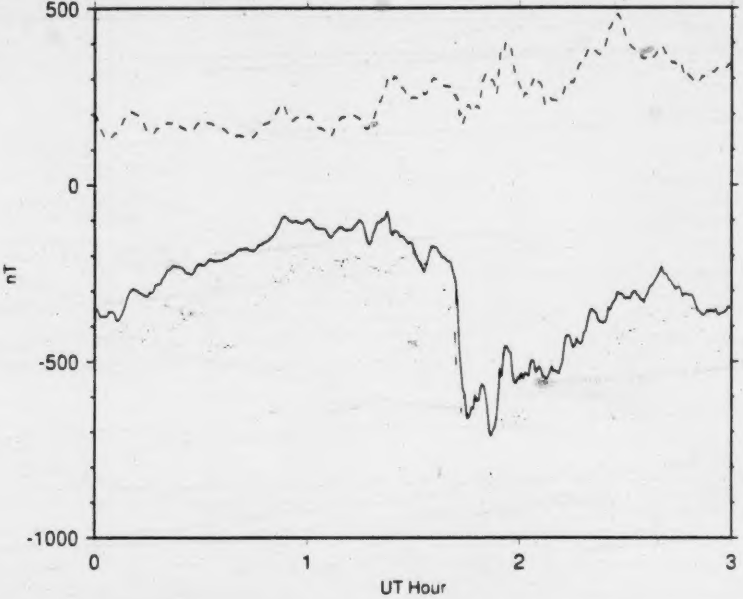


Figure 8.7 Electron flux at satellite SC 1986-019 on May 3 1986. Log scale with same energies as those of Figure 6.21.

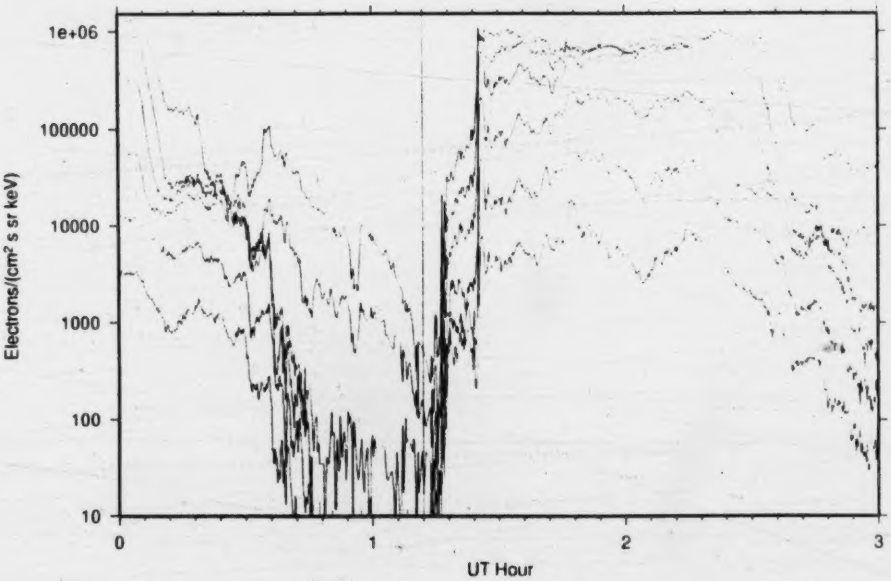
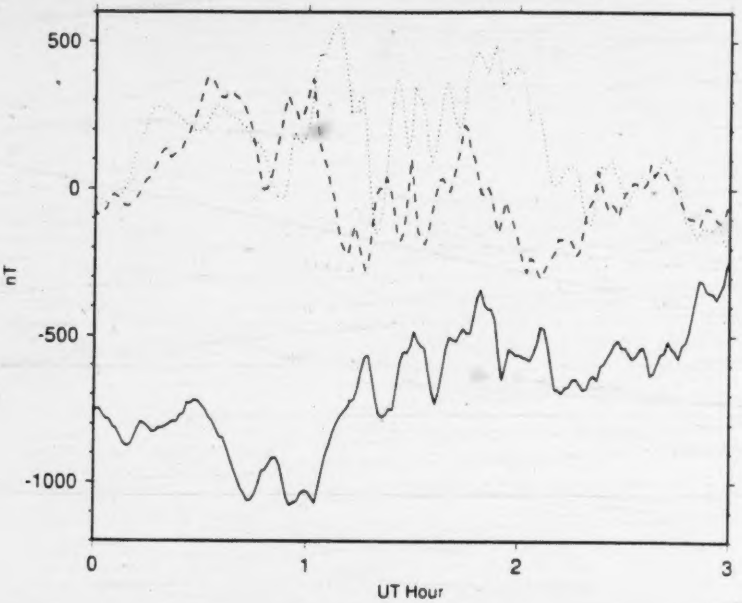


Figure 8.8 Centred dipole magnetic perturbations at Muonio on May 3 1986. X_M is indicated by a solid line, Y_M by a dashed line, and Z by a dotted line. Onset is marked by a vertical bar at 0112 UT.

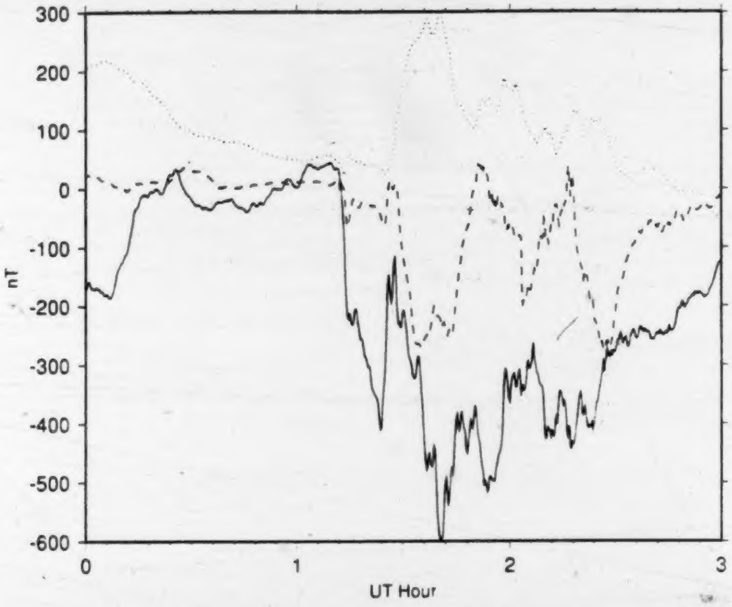


While the foregoing is not solid evidence that an onset did take place, it is suggestive. The X_M component declined by about 400 nT rapidly at Muonio, and declines were seen at other EISCAT stations also. The major onset which took place at 0111 is not strongly in evidence at these stations. Unfortunately, due to the presence of the Atlantic Ocean, there are few stations at the appropriate latitude to the west of the EISCAT array where more pronounced onset magnetic signatures might be expected. Stations in Iceland and Greenland are north of the electrojet at this time. At EISCAT, Ps 6 pulsations are present as indicated by quasiperiodic variations in the Y_M and Z components; they appear to be most coherent (see again Figure 8.9) immediately after the 0111 onset. Quantitative aspects of electrojet strength, including whether an onset took place, will be discussed in later section where the AFM method is applied.

Another aspect of this event is brought to light by Figure 8.7 in particular. The 1986-019 satellite, situated near the midnight meridian, detected strong and classic growth phase and onset signatures in the form of electron flux dropout followed by strong enhancement. The enhancement clearly associated with the 0111 onset, however, is not to pre-growth-phase levels but only to roughly one-tenth of those values. It is with another step at about 0125 that the 0 UT levels are once more attained. This is very suggestive of a second distinct onset. As noted above (see again Figure 8.1), the AE index rose only at 0130; this would be in response to the second onset as opposed to indicating the first one with a delay due to a station gap in the standard AE stations. That this is the case is shown with magnetic data in Figure 8.9. Lack of stations due to the ocean prevents use of northern hemisphere data from a suitable location, but the

Halley Bay Antarctic station is conjugate to a point near northern Newfoundland. The data presented in the figure have the local magnetic eastward (D or E) component and vertical component reversed to simulate observations at a northern hemisphere station.

Figure 8.9 Simulated data from the Halley Bay conjugate point on May 3 1986. Perturbations at Halley Bay are depicted with local magnetic D and Z components inverted. Vertical line at 01:12 marks first onset time.



At 0112 the classic onset signature of rapid decline in the northward (H) component is seen; that this is accompanied by little change in the Z component indicates that the causative current is nearly overhead. Halley Bay's conjugate point is bracketed on the south by Saint John's and on the north by Narssarssuaq (NAQ): at these stations Z, as seen in Figure 8.10, respectively decreases abruptly and increases, supporting the above contention. A second, not so precipitous decline in H at Halley Bay's conjugate point took place starting at about 0125, and was accompanied by an increase in Z, indicating current flow south of that point. This initially appears inconsistent with results at NAQ (see Figure 8.17), where no consistent Z deflection is noted at this time while X declines abruptly by 500 nT. The horizontal components in Figure 8.9 may usefully be compared with those (in Figure 8.4) from Poste de la Baleine, which have larger amplitude. The northward components are almost identical in shape, but this component is not very sensitive to electrojet location in latitude. The eastward components during the second onset are similar in shape but inverted, and only during this second onset are they strong. While the Z component at Halley Bay indicates that the first onset was at about its latitude, consistently with Viking images, it suggests also that the second onset was equatorward of it, while magnetograms from NAQ and PDB suggest an onset location poleward of the station. This problem is now addressed.

Figure 8.10 Z at Halley Bay's conjugate point, Saint John's and Narssarssuaq on May 3 1986. Solid line is the simulated Z component at Halley Bay, dashed the observed value at St. John's, and dotted that at NAQ. Time lines at 1:12 and 1:24 UT

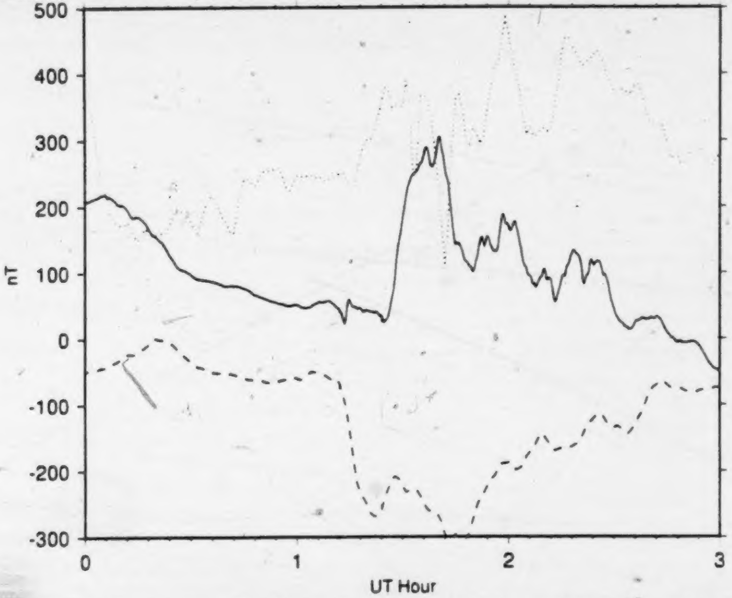
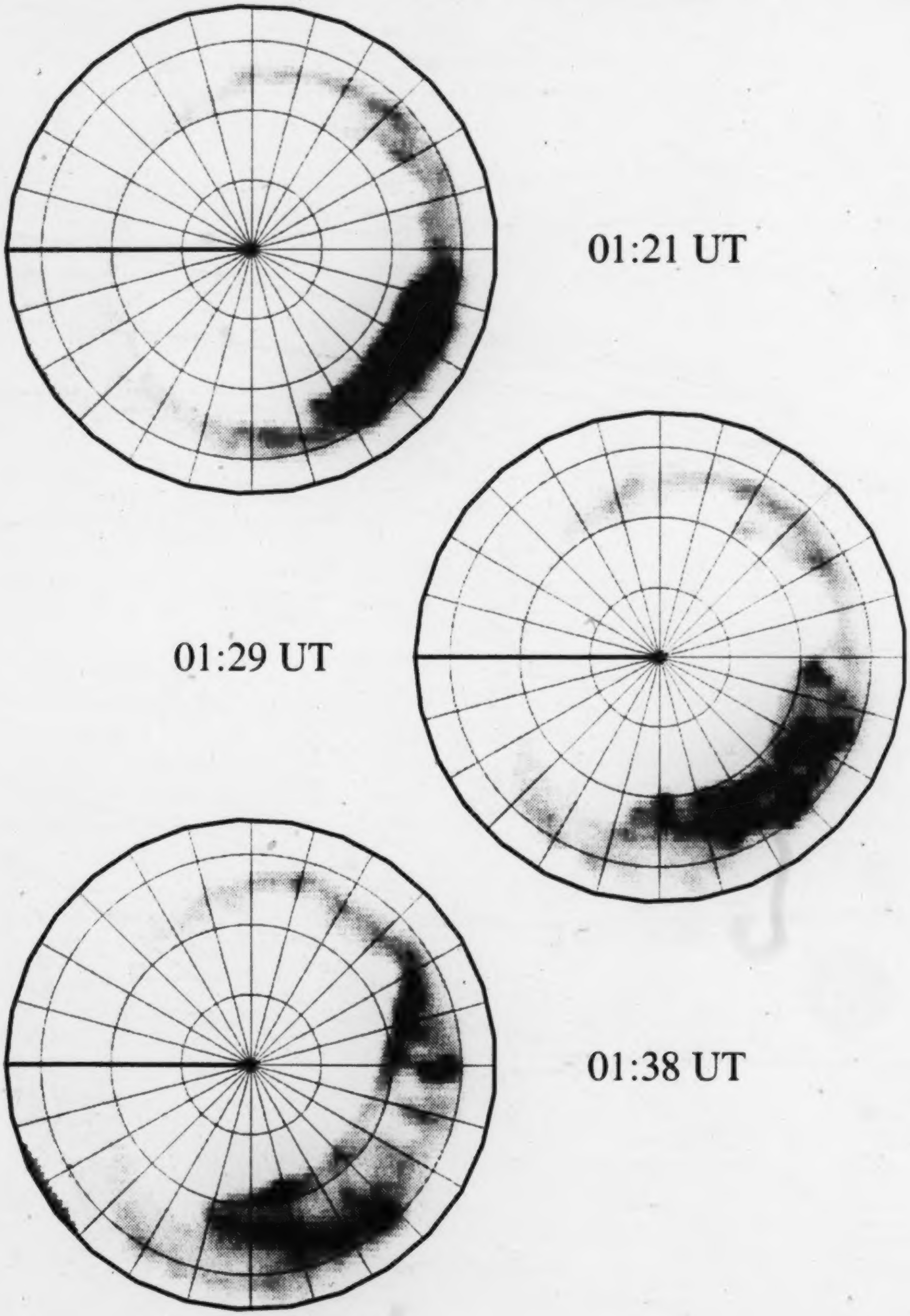


Figure 8.11 shows Dynamics Explorer 2 images of the southern hemisphere in an inverted grayscale. The noon local time meridian is to the left (and is shown by a heavy line) so that the orientation is roughly comparable to that in the Viking images of Figures 8.2 and 8.3. Much like in the post-onset image in Figure 8.3, the active regions of the late evening sector in the top image (for 0121 UT May 3) were very much brighter than any other features. In the Viking data, saturation had taken place: here the greyscale was adjusted for good contrast in other regions, effectively also saturating these regions. At 0121, the active region is below 70°, corresponding to what has been deduced above about the location of the electrojet. It is well south of the station NAQ which is at roughly 70° and at MLT 2300. PDB, at about 66° and MLT 1900, is near the poleward border of the active region at this time, a result consistent with the negative X and positive Z perturbations at this time shown in Figure 8.4. However, the -500 nT perturbation at PDB is not sufficient for it to dominate the AE index and so indicate an onset: morning sector perturbations from the westward electrojet do this (see Figure 8.8) as they are stronger. By 0129, the poleward border of the active region has expanded to above 70°. This accounts for NAQ having an onset-like drop in the local magnetic northward component (Figure 8.17) at this time. PDB is now close to the centre of the active region, which has complex structure, accounting for its strong negative X signature (it now dominates AE) and variable Z. The bottom image, from 0138, shows that a double auroral oval formed later in this event. Unfortunately, Viking imaging stopped around 0125, not allowing these later developments to be studied in detail. However, it is clear that a second onset, evidenced by an injection and Halley Bay magnetics, and simultaneous poleward expansion starting double oval formation, explain all aspects of the data.

8.11 Dynamics Explorer southern hemisphere images on May 3 1986. Images have been transformed to appear as if in the northern hemisphere poleward of 55°. Noon to left. Grid is at 1 hour local time intervals (15°) and 10° magnetic latitude.



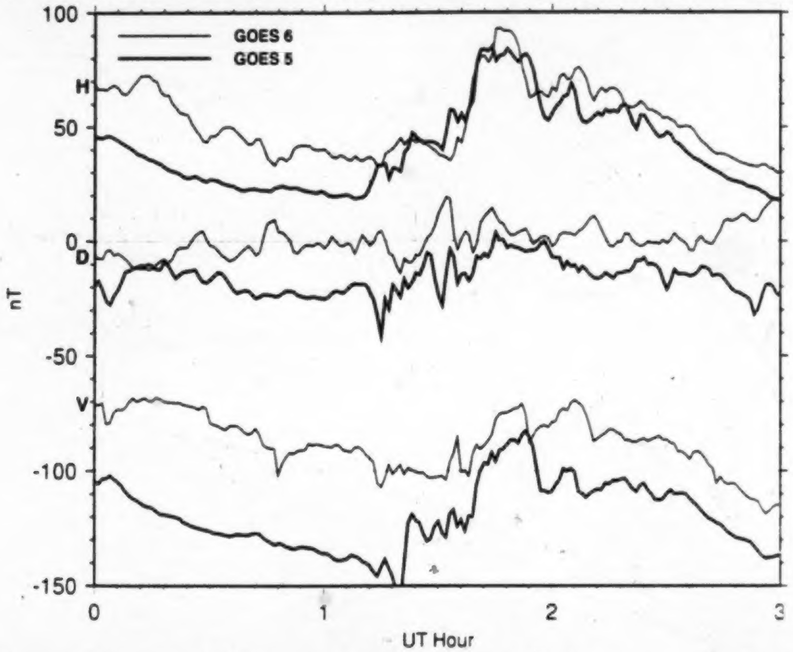
01:21 UT

01:29 UT

01:38 UT

A further insight into the onset activity may be had by looking at magnetic data from the GOES spacecraft. These are presented from the CDAW data file as Figure 8.12 (note if comparing to the figure for GOES 5 in Baker *et al.* [1993] that H here corresponds to Z there, D here to Y there, and V here to (negative) X there).

Figure 8.12 Magnetic field at GOES 5 and GOES 6 on May 3 1986. H is northward field, D eastward, and V radially outward, as indicated at side. Heavier line for GOES 5.



GOES 5 was above 75.4° W geographic latitude on this date, near the 0° centred dipole magnetic meridian shared by Ottawa and PDB, while GOES 6 was further west, at 107.8° W, roughly in the Churchill meridian. Both spacecraft would thus have subsatellite points at about 10° magnetic latitude and L values slightly greater than 6.6. Due to this northern hemisphere (magnetically) positioning of the spacecraft, the radial component is negative. At GOES 5, it decreased in magnitude after the 0111 nominal onset time, and at this time the northward component increased, consistent with a dipolarization. Shortly after this, however, the radial component's magnitude became larger, going to -157 nT (slightly off the graph) briefly at 0120. At about 0124 another dipolarization signature was seen in the northward and radial components. This is consistent with the presence of a second onset as inferred above. These signatures are weaker at GOES 6, and both spacecraft show a much larger dipolarization signature at about 0140. Before the 0111 onset, both spacecraft showed classic growth phase signatures. Another interesting aspect is the behaviour of the D component shortly after the 0111 and 0125 onsets: it shows opposite polarity perturbations at each spacecraft, with those at GOES 6 being positive, and those at GOES 5 being negative.

From the data examined so far, it is clear that the major onset at 0111 was not isolated. It was followed by at least one further onset, at 0125. There are indications that others took place also, most notably at about 0140. There is little reason to question the basic interpretation that a growth phase in the classical sense was underway from about 0000 UT to the time of the first major onset. It seems likely that a small onset took place during this period in the midnight sector, but relevant ground data were available only from a small region where the westward electrojet was already very strong, making interpretation difficult. Features in Viking images were discussed (see again Figure 8.3) which are located poleward of the main auroral oval and in the midnight sector. These features are somewhat reminiscent of the late phases of the development of the 'eye' of the substorm as discussed in Chapter 7 and in Rostoker *et al.* [1987]. This also suggests that an onset at roughly 0030 did occur. If such an onset did take place, a relevant question would be to ask what limited its growth. In this sense, the 0030 occurrence might be regarded as a 'pseudobreakup'.

It has also been noted that Ps 6 pulsations, as observed in the morning sector in Scandinavia, strengthened at the time of the 0111 onset, which took place well into the evening sector. It has usually been claimed that such pulsations are a property of the recovery phase of substorms, which would imply that they do not immediately follow the onset (which starts the expansive phase). It is tempting to ask if this conclusion has not been biased by use of single meridian chains, where one would have to wait a considerable period of time after a midnight sector onset before morning sector features such as Ps 6 become observable. The fact that observations in a single meridian often show onset-like behaviour, later followed by Ps 6 activity, could well be an artifact of the Earth's rotation. It may be more correct to regard Ps 6 as a feature of expansion or even growth phase. Connors and Rostoker [1993] found them to be present before the 0111 onset, that is to say, during the growth phase of this event, and seemingly little affected by that onset.

Due to the questions raised by preceding studies, this introductory section has been rather detailed. In the next section an overview is given, in part based on what has been discussed here. We then proceed to modelling, both on a local and global scale, to attempt to further clarify some of the points raised above and to get a global view of this event.

b. Chronological Summary

To assist in evaluating the most useful aspects of this event sequence to study with Automated Forward Modelling, a brief summary will be given as for other events. Despite their potential importance to the overall analysis, events outside of the CDAW 9C period, 0000 to 1200 May 3 1986, will be discussed only briefly. As noted above, the activity level immediately preceding the event was high. A large (-1500 nT) bay signature observed at Leirvogur (LRV) at about 2320 on May 2 fits the trend noted by Baker *et al.* [1993] for recurrence of major events every 2 to three hours (the others they noted being on May 3 at 0130, 0430, 0530, 0715, and 0920). This bay undoubtedly corresponds to the 2330 AE excursion noted above, but does appear as measured in one-minute data to be slightly larger in magnitude than AE would have indicated. In the table below, occurrences identified from the wide-ranging paper of Baker *et al.* [1993] are denoted with (B). The onset time of the 0111 onset is uncertain by about one minute.

Table 8.1 CDAW 9C Onset on May 2 and 3 1986

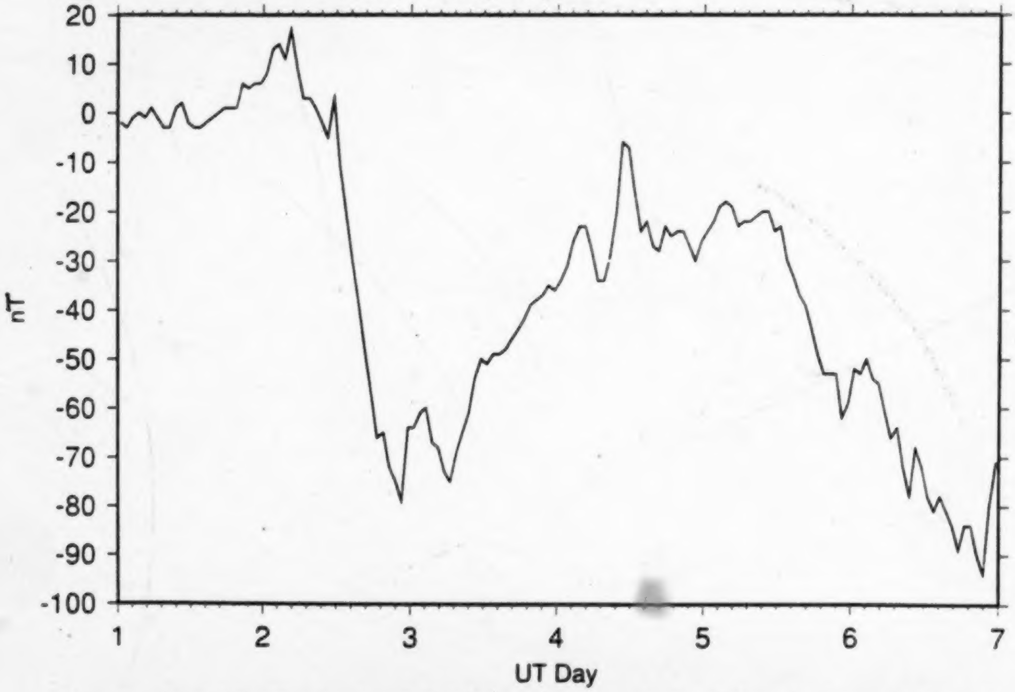
Time (UT)	Activity	Observational Basis
May 2		
2315	major onset (to -1500 nT)	LRV analogue recording
May 3		
0000	active auroral oval and strong electrojets	AE=1000, DE-1 (B)
0000	beginning of growth phase	GOES-5 (B), 1986-019 (B), SCATHA particles and fields (B)
0030	minor onset	EISCAT magnetics, particle injections, possibly later Viking images
0111	major onset	multiple sources (B)
0112	optical onset	Viking images
0127	second onset (AE onset in (B))	SC 1986-019 electrons, ground magnetics, GOES
0140	possible third onset	GOES, ground magnetics

It will be noted that the onset identified as the AE onset in Baker *et al.* [1993] is here identified as a separate onset for the reasons given in the preceding section.

c. Modelling Results

The majority of the data used was taken from the CDAW 9 CD-ROM and the USGS 1985-1989 CD-ROM described in Chapter 6. Most data were prepared by box-averaging with a 60 time point window (one hour for one minute data, 10 minutes for 10 second pulsation data), subtracting the smoothed value at 1100 UT, and rotating if necessary. The listing for station Meanook had the last five values (near 1200 UT) missing and these were replaced with the values from 1154 UT. Values at Murmansk (supplied as perturbations) were missing for 4 minutes prior to 01 UT and at 0114. these were filled with linearly interpolated values. Data were not available for Sitka on either CD-ROM, which is unfortunate since this is one of few stations showing evidence of the evening sector eastward electrojet. For the lower latitude stations it was frequently noted that the simple subtraction of the smoothed 1100 UT May 3 value left a seeming bias in the perturbations at times of lower activity. For most of these stations, the May 1 quiet day values were available and data were prepared by subtraction of the smoothed quiet day from the May 3 values. This still left a negative bias at the end of the day, so Dst values were obtained and are shown in Figure 8.13. The Dst values at the end of May 3 are of the same order as the offsets noted in the low latitude data and it is concluded that these depressed values were a ring current effect. Since this effect is incorporated into the global model, the data values as produced by quiet day subtraction were used with no further adjustment. For electrojet modelling, these stations were not used, and the D_{st}-like parameter was generally omitted.

Figure 8.13 Dst index values for early May 1986.

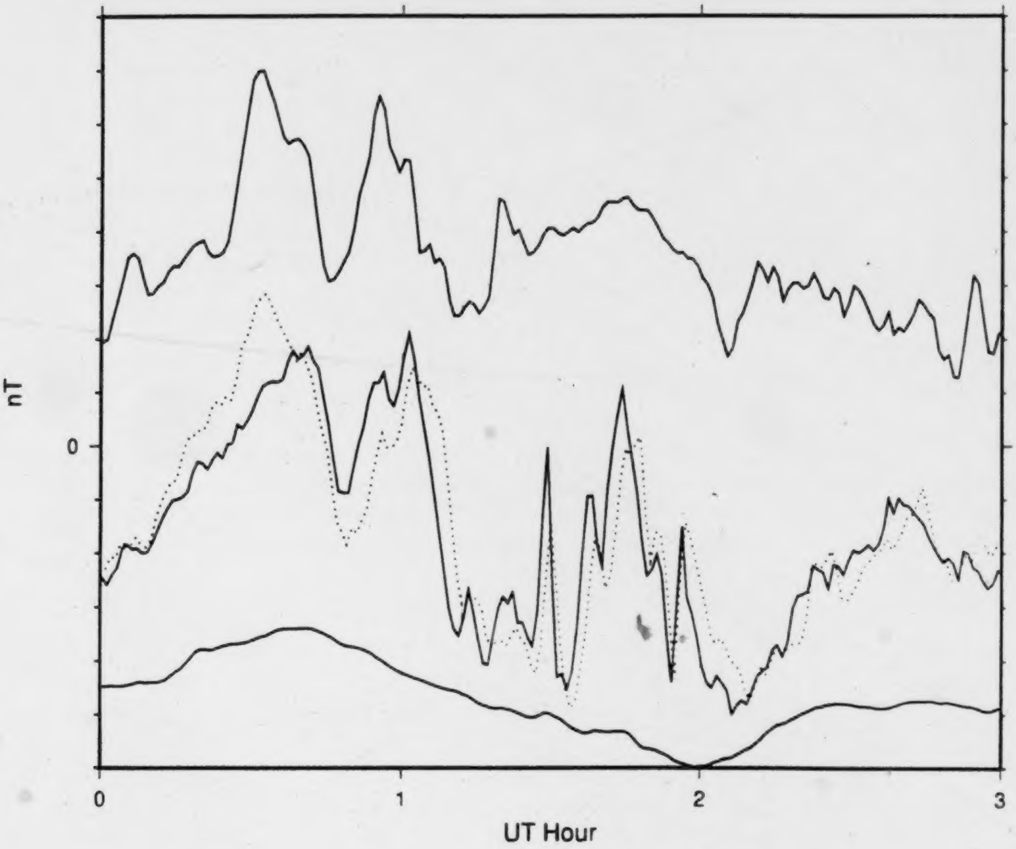


Data for Lerwick, Hartland and Eskdalemuir were obtained directly from the British Geological Survey [Toby Clark, personal communication, 1993] and transformed into the centred dipole system. Greenland perturbation data were obtained from the Danish Meteorological Institute [Eigil Friis-Christensen, personal communication, 1996] for the May 1 to 3 period. These were prepared using subtraction of the smoothed quiet day May 1 and compared to values taken from the CD-ROM (and already prepared by subtraction of the smoothed 11 UT value) for some stations (NAQ, GDH, THL). The comparison was very good, and data prepared by quiet day subtraction were used for the entirety of the Greenland data. Greenland data for modelling were also prepared for May 2. Baselines from May 1 for station DNB were bad very late in the day, so for the May 2 modelling the raw perturbation file from DNB was used. As modelling only for early in the UT day was envisaged for May 3, the DNB file prepared by quiet day subtraction, with bad data at the end of the day, was retained. After some preliminary global modelling, perturbation values from the early CANOPUS stations Gillam and Eskimo Point were added, bracketing Churchill. Data from these stations were not available before 0025 UT. Local magnetic values from Iqaluit were also added, as taken from the CDAW CD-ROM and baselined relative to 11 UT. X and Y values at Nurmijarvi were supplied by Ari Viljanen [personal communication, 1997].

In attempting to understand the electrojet behaviour, the presence of features which cannot be represented by a simple electrojet model is a hindrance. In this case, the morning sector westward electrojet has current systems associated with it which produce Ps 6 pulsations, as noted by Connors and Rostoker [1993]. The magnetic signature of these pulsations is shown in Figure 8.14. Such pulsations correspond to the perturbations from an eastward drifting equatorward ionospheric current system with its corresponding field-aligned currents [Kawasaki and Rostoker, 1979]. This is seen in the diagram through the fact that the coherence between pulsations observed at Pello and those observed at Sodankyla, roughly 125 km further east, is high, with the patterns appearing at Sodankyla somewhat later than at Pello. The perturbations observed at these stations cannot be adequately represented by a simple electrojet model consisting of westward flowing current. As would be expected, the Z component is also affected by the overhead passage of current and does not follow the simple pattern associated with a pure westward electrojet. The northward component, though, is little affected by a north-south aligned ionospheric current flow. A complex electrojet model, featuring north-south as well as east-west currents, is in principle needed if the morning sector electrojet is to be modelled. However, it was found that the low coherency of these particular pulsations between stations allowed reasonable modelling results to be obtained with a simple electrojet model of large scale size. In this case the lack of coherency across a wide region would make the Ps 6 pulsations appear as noise, which may have affected speed of convergence but which appears not to have dramatically affected overall results.

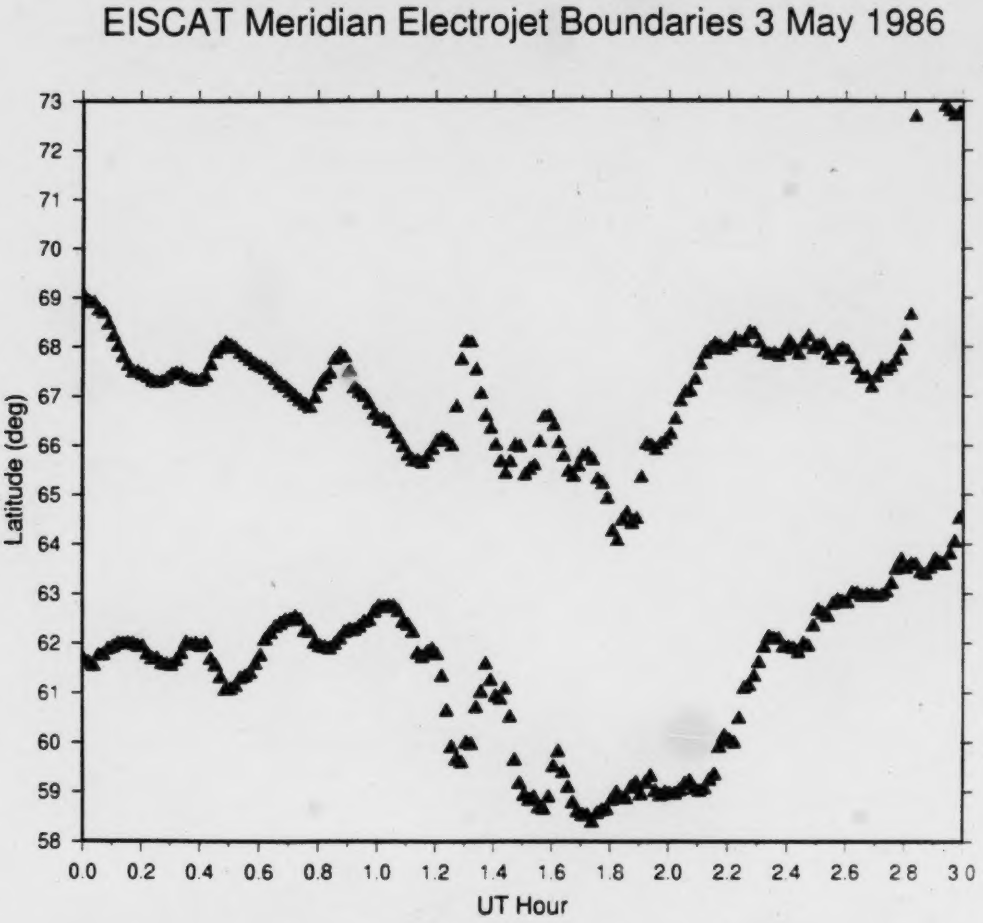
It is clear from the perturbations shown in Figure 8.14 that the Ps 6 signal is clearest near the southern border of the electrojet, which was near the Pello/Sodankyla latitude. This result is consistent with the findings of Connors and Rostoker [1993] based on UV images in which the aurorae associated with the electrojet may be seen. In that study, spots of emission, which correspond to the upward field-aligned current of the Ps 6 current system, were located in the images near the equatorward border of the electrojet (i.e. of the auroral oval).

Figure 8.14 Ps 6 pulsations in the EISCAT meridian on May 3, 1986. Curves have been displaced vertically for clarity and the vertical axis has ticks at 100 nT intervals. The Y_M component is shown for stations Soroya (top), Pello (middle), Sodankyla (middle dotted line), and Nurmijarvi (bottom).



As a first test of the possible detrimental effects of the presence of Ps 6 pulsations, it was decided to proceed with a simple electrojet model. The Y_M component, which responds to the Ps 6 pulsations, is given a very high error value and effectively removed from consideration. In previous chapters, this approach has been used even though Ps 6 pulsations were not present and is simply due to the fact that the major perturbations arising from a long electrojet with current flow parallel to lines of latitude are in the X_M and Z components. The Z component, which shows the effects of Ps 6, must be used in electrojet modelling since it is also a sensitive indicator of the electrojet boundaries. By using enough stations spread out east-west, the relatively small longitudinal scale of the Ps 6 currents can be taken advantage of. Only in one specific meridian at a time will the Z perturbations be seen, and the average Z perturbation over a longitudinally distributed region will be that of the electrojet Modelling done in this way produced the electrojet boundaries shown in Figure 8.15.

Figure 8.15 Electrojet boundaries in the EISCAT meridian on May 3 1986.

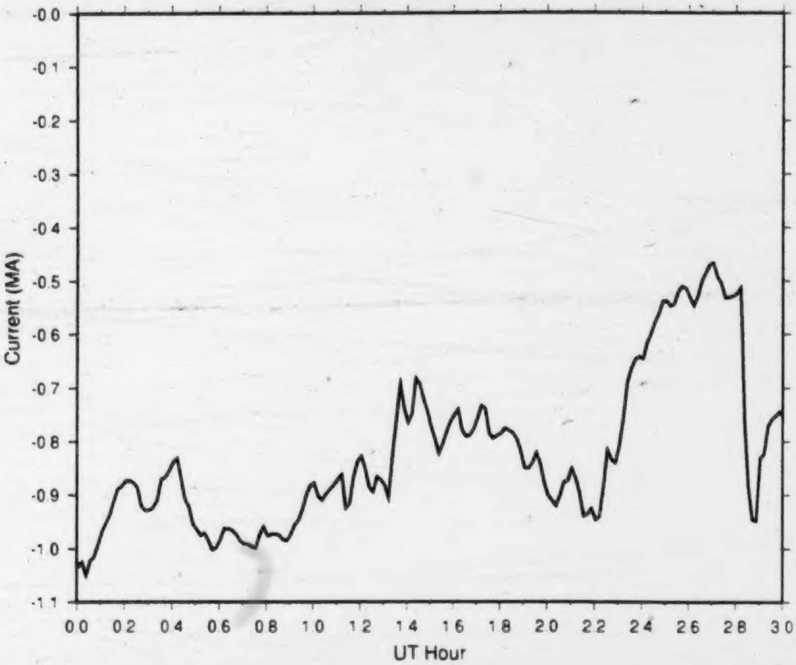


In interpreting this diagram, it is important to consider the latitudes of contributing stations. These ranged from Nurmijarvi in the south, at magnetic latitude 57.7°, through Soroya in the north at 67.2°. No Z component data were available for Nurmijarvi, and there is a gap to the next station at Pello at 63.7°, after which several stations are available, including Murmansk east of the immediate EISCAT region. It is thus clear that more weight should be given to a result indicating a boundary in the range of about 62° to 68° than to one outside this range. Values south of this range would have higher weight than those to the north due to the presence of at least some data from Nurmijarvi. With this in mind, one finds relatively stable and firm estimates of the electrojet boundary in this sector (which is near midnight) between 0 and 1 UT.

There is some indication of widening of the electrojet at about 0130, which supports the contention that a substorm intensification took place then. There is, however, no indication of a narrowing of the electrojet prior to this putative onset. There does appear to be narrowing prior to the undisputed 0111 onset, followed by widening both to the north and to the south. The narrowing and widening cycle is repeated for the 0130 onset, and there appears to be general equatorward motion in this latter half of the UT hour. Inspection of the available X_M component at Nurmijarvi supports this contention in that a deepening of the existing negative bay appears when the routine has indicated southward motion of the southward electrojet border. After 0200, there appears to be rapid northward motion of the overall auroral oval. Toward the end of this UT hour, a rapid poleward expansion is indicated. Since the poleward border is found by the routine to be well north of the northernmost station, little weight can be given to this result, and inspection of data from further north would be required to find out what actually took place at this time. That the southern border is located in the middle of the array and moved northward does suggest that northern motion did in fact take place, but undoubtedly not to the extent indicated by the northern border found by the modelling routine. Comparison of the times of border excursions with the times of Ps 6 pulsations seen in Figure 8.14 shows that there is little relation between the two. The Ps 6 signal does not appear to have interfered with the ability of the routine to find electrojet borders. One may speculate that if the electrojet signal could be simulated by running the routine and specifying its parameters as already found here, then the superposed Ps 6 signal could be modelled. This was not attempted.

The discussion of the westward electrojet may continue by examining the current found in the simple model. This parameter is shown in Figure 8.16. The eastward current is large in magnitude and negative, as emphasized by the plot vertical axis including 0. The strength of the morning electrojet is typical of very active times. After decreasing in magnitude until roughly 0.4 UT (0124), this current has about a 25% intensification. This is near the time at which an onset is claimed to have taken place. Such an onset might not normally be noticed against the strong background current of the electrojet, but this is another indication that it did occur. The undisputed onset at 0111 does not have any signature in the electrojet current. The upward current at least of the onset was located well into the evening sector (recall Figure 8.3 and see discussion on global modelling below). After the onset, the current in the westward electrojet declined in magnitude. This is reflected in the AE index (Figure 8.1) as has been pointed out by G. Rostoker. At about 0130 the second onset took place and at that time the responsible station for the AE index shifted from Scandinavia to Poste de la Baleine, Quebec. There was only a small response to this second onset in the morning sector current strength. After 0200, the current in the westward electrojet declined rapidly in magnitude, corresponding to the northward motion of the overall auroral oval. Much as the rapid poleward motion of the northern border shortly before 0300 UT is spurious as discussed above, so also is the corresponding rise in the magnitude of the electrojet current. The spurious widening has required an increase in current in the model to maintain about the same perturbations despite the current being forced to spread out over a larger latitudinal range.

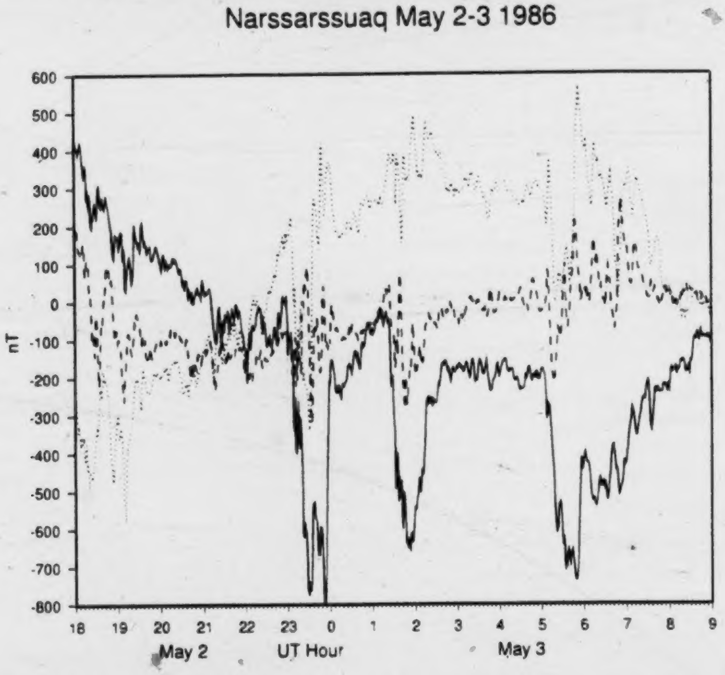
Figure 8.16 Eastward current over EISCAT Meridian on May 3 1986. Negative values indicate westward current flow typical of the morning sector electrojet.



The basic conclusions about the morning/midnight sector westward electrojet in this case are that some signs of an onset are present at about 0030, and that latitudinal motions associated with onsets are observed at 0111 and about 0130 with little increase in current magnitude. After about 0200 there is poleward motion of the auroral oval in this sector accompanied by a decrease by almost a factor of 2 in total current. Before about 0300 there is indication of further poleward motion but by this time the poleward border has moved into a zone of no data and no reliable inversion is possible.

Before proceeding to global modelling for the period early on May 3, data from Greenland for late in the May 2 UT day will be examined in an attempt to study the beginnings of the substorm sequence. It has been mentioned above that analogue data from Leirvogur were examined for this pre-CDAW time period and that a large onset took place at about 2315. A magnetogram covering this time period for Narssarssuaq (NAQ) is shown in Figure 8.17.

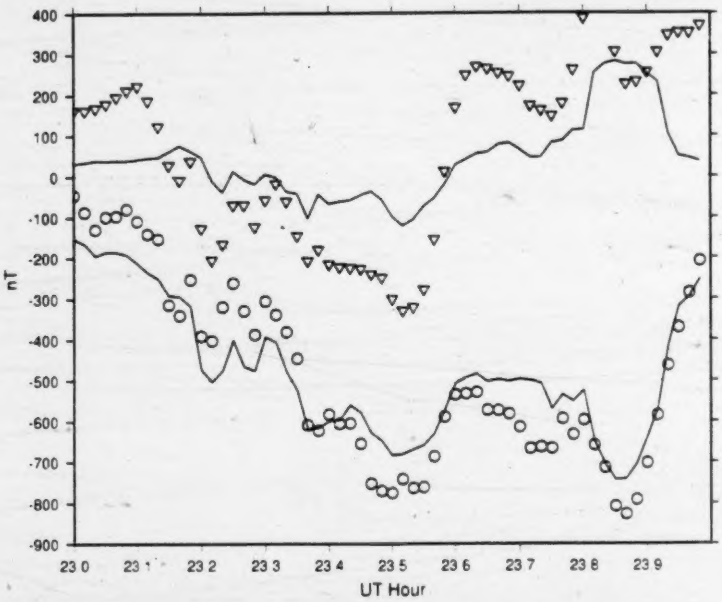
Figure 8.17 Magnetogram for Narssarssuaq spanning May 2 to May 3 1986. Perturbations are in local magnetic coordinates with H solid, E dashed, Z dotted.



It is immediately obvious that the onset at about 2315 May 2 as referred to above was not the first onset during the immediate pre-CDAW period. A sharp decrease by 100 nT in the H (local magnetic north) component almost exactly at 2300 preceded the major onset by about 15 minutes. Over 200 nT of additional decrease took place at 2305, and the Z component suddenly going negative at this time suggests that the onset location was to the north of NAQ. This remained the case at the time of yet further decline corresponding to the 2315 onset. At 2330, Z abruptly became positive once more, and the H decreased, indicating another onset, but this time equatorward of NAQ. The general situation of positive Z after this time, at this most southerly Greenland station, suggests that electrojet was concentrated south of Greenland. This general positioning of the electrojet in this sector appears to have been maintained until another major onset at about 5 UT. Although very distinct at NAQ, this onset is not very distinct in AL (see Figure 8.1), but due to an accompanying response in AU, the AE index shows it quite clearly.

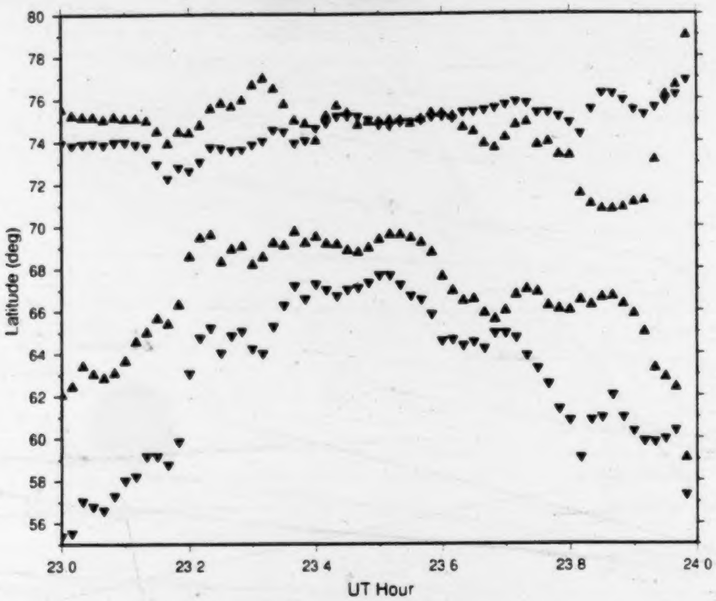
Returning to discussion of pre-CDAW conditions, modelling was done using the Greenland stations alone. Lack of other available data restricted the modelling to them. Two initial conditions were used, one a large current case (1 MA), and the other a small current case (0.05 MA). It was noted that the large current case at times had solutions with a very wide electrojet as end states, while for the same times the small current case had narrow electrojets. At other times the two results were in good agreement. The discrepant results between the two starting conditions have been examined before: if there is a lack of constraint from adequate latitudinal coverage, a large current with a wide electrojet covering regions where there are no stations can give a good fit where there are stations. One cannot truly determine the electrojet width in such a situation.

Figure 8.18 Observed and modelled perturbations at Narssarssuaq on May 2 1986. Data are shown by discrete symbols: inverted triangles for Z and circles for H. Adjacent lines are model results, not always showing a good convergence.



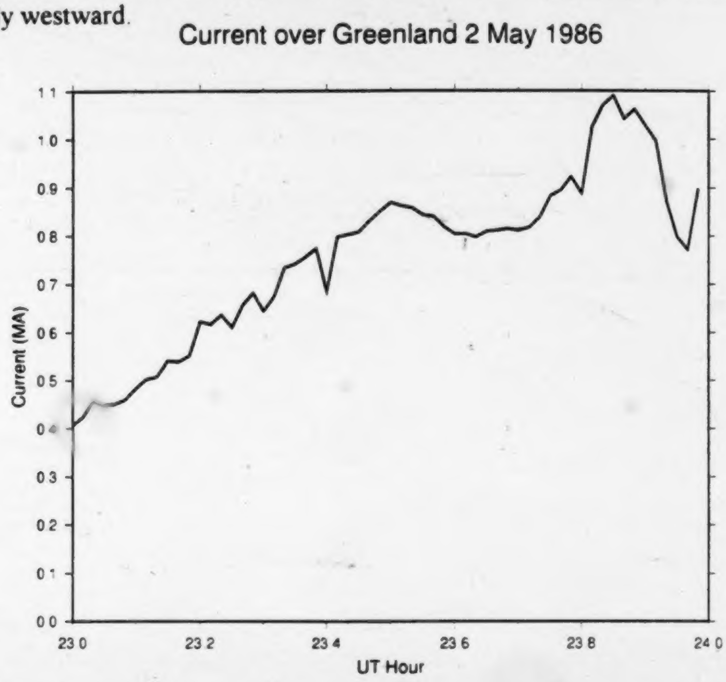
To determine what is reliable in the modelling results, it is important to examine how well the input data matched the model, that is, the fit of the magnetograms. The modelled local magnetic H and Z components for NAQ (CD latitude 70.69°), as found from the large initial current model, are shown in Figure 8.18. It may be noted that the modelling does not at all times give a good fit to the data. In general also, it was noted that the times of discord between the small initial current and large initial current convergences, as noted above, were those when the modelling results gave a poor match to the initial data. As this station is the most southerly available, a positive Z value here generally indicates that the electrojet centre was south of the array of stations. It will be noted that those are the times at which convergence was poorest. At all times with a negative Z value, so that the centre of the electrojet was at least somewhat within the array, reproduction of the data values by the model was better. This was also the case at the next two stations further north, but the model values yet further north (starting at Sondre Stromfjord at 76.93°) were much smaller than the observed values. This suggests that a higher latitude current system was present in addition to the one centred roughly at 70°.

Figure 8.19 Electrojet boundaries in Greenland on May 2 1986. Latitudes at west end of electrojet are shown as triangles, at east end as inverted triangles.



With the knowledge that the convergence was not good at certain times, an informed look at the electrojet boundaries may be taken. These are shown, for the large in Figure 8.19. Since there was a possibility of electrojet tilt, and since this is regional modelling, not a meridian inversion, differing latitudes at either end (fixed longitudes of 15° and 105° in centred dipole system) of the current system were allowed. The local magnetic eastward (E) component was also included in the modelling: it was generally neither large nor well modelled. It is seen that during the times of best convergence the latitudes do not vary greatly between the ends of the system. Although the inferred border values are not reliable, it also appears, consistently with what the Z component of the magnetogram would suggest, that the electrojet was further south near the beginning and end of the UT hour modelled. This in turn suggests that the electrojet was further north during the onset. Finally, the current in the regional model for the pre-CDAW electrojet is shown in Figure 8.20. The results shown here must be interpreted with great caution. Before about 23.2 UT, the electrojet is known to be too broad in the model, and a larger current is used by the model to make up for this error. Thus values before 23.2 UT are likely less than the 0.4 to 0.6 MA indicated in the figure. How much less is a matter of speculation: this is a time of active electrojet and the figure found is not unreasonable during such a period. Nonetheless the data presented here do not allow the statement to be made that that value came from the model. The 0.6 to 0.8 MA during the active period from 23.2 to 23.8 UT does seem well founded: also in light of the perturbation at Leirvogur seen in analogue data and reflected in the AL index, this does not seem an unreasonable value. The rise after 23.8 UT to values over 1 MA appears once more to be related to widening of the electrojet, and since much of the widening indicated is to the south, where there are no stations, it is uncertain to what extent this is real. The falloff in the H magnitude and rise in Z at Narssarssuaq (see again Figure 8.18) may reflect equatorward motion of the electrojet or decrease in current or both. Since the model did not represent these values well, it does not really tell us.

Figure 8.20 Current over Greenland on May 2 1986. The current flow direction is primarily westward.



To complete discussion of pre-CDAW conditions, some discussion of the observations at more northerly Greenland stations is in order. It will be recalled that the simple electrojet modelling attempted above did not account for the magnitude of perturbations observed north of about 75°. The most useful way to analyze this discrepancy is through examination of latitude profiles. Such profiles from relatively late in the event are presented in Figures 8.21 and 8.22. It can be seen from the appearance of the profiles that the electrojet has been modelled with roughly half of its expected signature present and the implications of this for accuracy of modelling south of the range of stations has already been discussed. However, even in the more northerly portions the simple electrojet model, although the optimal match to the data, clearly does not fully represent it.

Figure 8.21 Latitude profile of H and Z components in western Greenland at 2336 UT May 2 1986. Observed local magnetic H as circles, Z as triangles. Nearby lines are model result. Grey bar indicates on-graph portion of model electrojet.

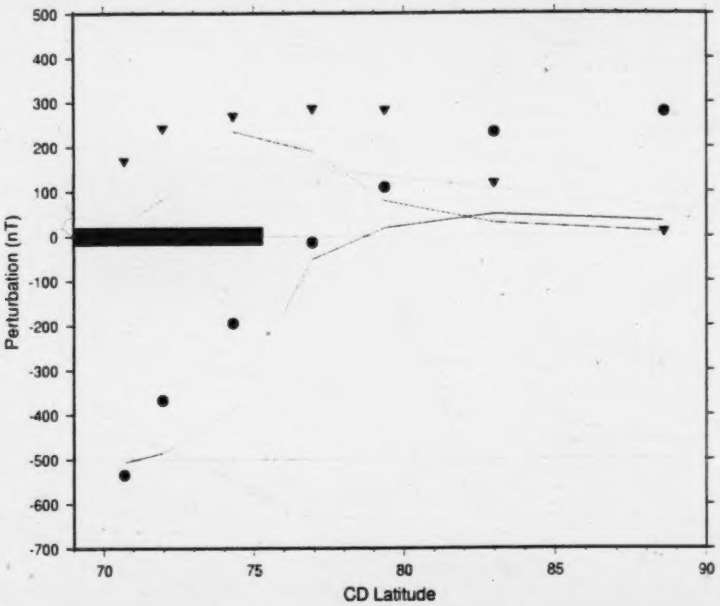
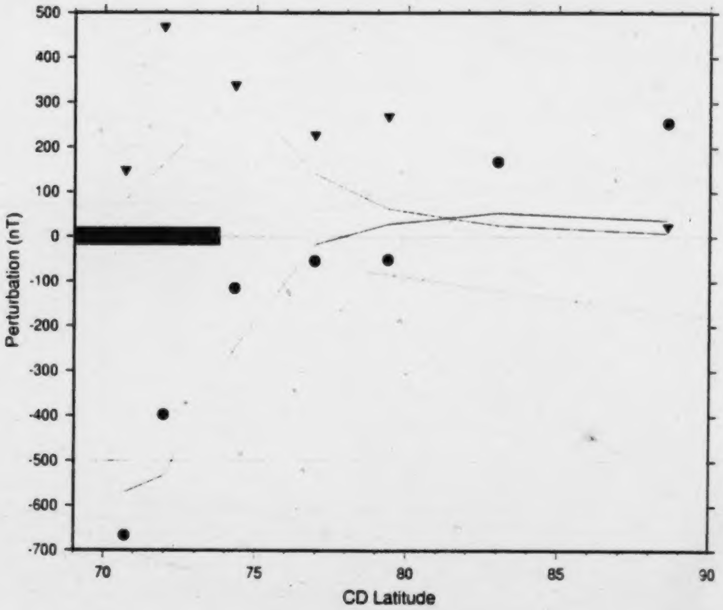


Figure 8.22 Latitude profile of H and Z components in western Greenland at 2345 UT May 2 1986. Data keyed as in previous diagram. H and Z points overlap at 83°.



At time 2336 (23.60) UT the H and Z drop off to the south in a manner appropriate to a rather wide electrojet. Even then, there seems to be excess positive Z in the 77° to 80° range. This could conceivably be an edge effect of what appears to be a very high latitude eastward flowing current hinted at by positive H values north of 80°. By 2345 (23.75) UT, the southerly electrojet current appears to have strengthened and narrowed as indicated by both the H and Z components. The optimization routine has been affected by the positive Z to the north and prevented from sufficiently narrowing the electrojet, although it has been narrowed slightly. The peak in Z has not been matched by the routine due to this biasing. The H component has decreased at 80° and due to a decrease in Z at 77°, the nearly unchanged Z there makes the Z profile appear bimodal. These changes suggest that the northern border of the electrojet has an intensified current flowing in it, a finding consistent with earlier results. This is also consistent with the later presence of high latitude features over Greenland as discussed above in the context of Viking imaging. By that time (0050 UT May 3) there appears to have been appreciable southward motion of the overall electrojet to roughly 60° (see Figure 8.2) if it is indeed the main electrojet current which has been detected here at about 70°.

To explore this situation further, the Greenland electrojet was modelled for May 3 during the time period 0 to 3 UT, and the results briefly described here. Since it was known that the electrojet expanded southward at the time of the 0111 onset, the station St. John's was included (centred dipole latitude 58°) in the modelling. Despite the large gap of 12° between it and NAQ, this inclusion ties down the southern border of the electrojet. Greenland stations north of 80° were not used for reasons of speed and because they appear to contain little information about the electrojet itself. An initial attempt using the initial small current conditions as described above led to very poor results except at the time after the onset at about 0125 UT as observed at NAQ. This was verified by the small departure from initial conditions as well as by reconstructing a NAQ magnetogram. This run was then repeated from the initial large current condition, with 1 MA of electrojet current. In this case, results up until about 0130 UT were quite good in terms of a good fit at St. John's and Greenland stations. Before 0111, the electrojet borders consistently came out to be about 64° to 71° with a current of about 0.7 MA (westward). Comparison with Figure 8.10 would suggest that the equatorward border is not properly indicated by this result: Halley Bay, with conjugate point at about 66°, does not show the perturbations in H or Z which would be expected for being near the border of the electrojet. During this time period, the perturbations at St. John's are close to zero. These would not change much even with relatively large changes in the position of the relatively distant electrojet. Thus inclusion of St. John's is in fact not very useful in improving a bad layout of stations. It is likely that the electrojet in the west Greenland sector extended from about 72° to 69° and was slightly weaker than the inversion showed. After 0111, the usefulness of St. John's improves considerably since non-negligible perturbations are present in the magnetic records. The inversion indicates a southern electrojet boundary of about 59° while the current through the meridian rose to about 1 MA. This result is not inconsistent with what is observed at Halley Bay. It is an interesting result since it indicates that the onset in the near-midnight sector was equatorward of the pre-existing electrojet in that sector. This appears to be the situation inferred by Baker *et al.* [1993] from considerations of luminosity in the evening sector, but here it is magnetic data which suggest it. During this 0111 to 0130 period, the southerly onset is firmly supported by the data, much as discussed in the introductory section. After about 0130, a period of poor convergence occurred, with St. John's data in particular poorly matched even when the station weighting was increased to twice that of the other stations. The reason primarily appears to be a larger than expected Z perturbation in central Greenland, although overall southerly motion of the electrojet may also play a role. This situation is illustrated through a latitude profile in Figure 8.23. It is seen that the Z component is large and near-constant over a large latitude range. This feature is not encompassed by the simple electrojet model in use and indicates an active poleward border after substorm expansive phase onset (likely the second onset at 0125). This conclusion is entirely in agreement with what was inferred using DE images (Figure 8.11) in the introductory section. Figure 8.23 also illustrates the difficulty of interpretation since the latitude range of available stations clearly does not offer much coverage of the electrojet. For this reason no further attempts were made to use electrojet modelling. In a situation like this the inadequate station coverage limits the reliability of such modelling.

Global modelling was next attempted and is described below.

Figure 8.23 Perturbations and model for western Greenland at 2.6 UT on May 3 1986. Caption as in previous figure; latitude scale identical although northerly stations not included.

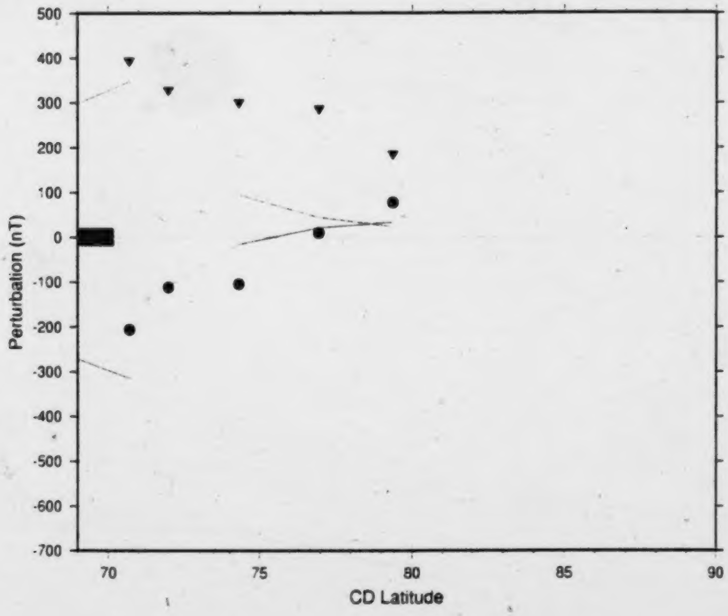


Figure 8.24 shows the stations used in modelling on the same projection as that used for results in figures below. This projection does not show results in the Scandinavian sector well (Nurmijarvi is in fact slightly outside of the region shown) but does have the advantage of giving a good view of the region of primary interest which is eastern North America. In order to at least show the Scandinavian region while retaining a portrait mode map (long axis vertical), the sunward direction is generally toward the left rather than towards the top as in other chapters. The sunward direction is shown by a symbol plotted at 62 degrees in the figures which follow. The midnight sector is to the right, and much of the region of interest is in the early evening sector, toward the bottom in the maps.

In the usual manner, an initial condition was set up based on preliminary inspections, and modelling runs based on that were done. The current systems consisted of an evening sector eastward electrojet, a morning sector westward electrojet, and a 'current wedge system' running northward into the evening sector, north of the eastward electrojet. The morning sector electrojet was tied to the wedge by constraining the field aligned currents near their intersection to coincide. Since the period of most interest fell in the timespan of 1 to 2 UT, only this period was modelled. Inspection of magnetograms showing data and model results indicated that poor convergence occurred in any timestep past 1.7 UT. We will focus discussion on the events at 0111 and 0130 UT. The model runs help in interpreting these early CDAW events, while they do leave room for improvement. Snapshots from these global modelling runs will be discussed now.

Figure 8.25 shows the conditions at 0103, just prior to the 0111 onset. The westward electrojet in the model has 0.62 MA of current and crosses the EISCAT stations accurately, with a reversal in the sign of the Z component. The magnitude of the current from global modelling was found to be smaller by about 30% than that shown in Figure 8.16. The lengths of perturbation arrows are not apparent in Figure 8.25, but it was verified in the model printout that modelled values were too small. In Iceland the perturbations due to the model were also too small, while at NAQ they were very good, and the rest of the western Greenland stations show positive X_M perturbations which likely do not have to do with the electrojet. The system reserved to be the current wedge contains westward current flow (but only .04 MA) and rises into the evening sector to high latitudes. At Mould Bay and Baker Lake the general sense of the model perturbations was as observed but the magnitude was smaller. At Cambridge Bay there was an eastward component to the magnetic perturbation which the model does not reproduce. Most importantly, the stations near Churchill show a primarily westward perturbation. This effect has been noted in other events. Its significance is not yet understood but it appears to be characteristic of the pre-onset phase. However, the most important aspect brought out through this model fit is the strength of the evening sector eastward electrojet. From accurately modelled perturbations at College, Barrow, Yellowknife, and Glenlea, it appears to be correctly placed, and the modelling shows a total current magnitude of 0.39 MA at this time.

Figure 8.24 Locations of stations used in modelling on May 2-3 1986. Dots not having a station code were in the database but without accurate data available during the period of interest. -40 degree magnetic median is at centre of map.

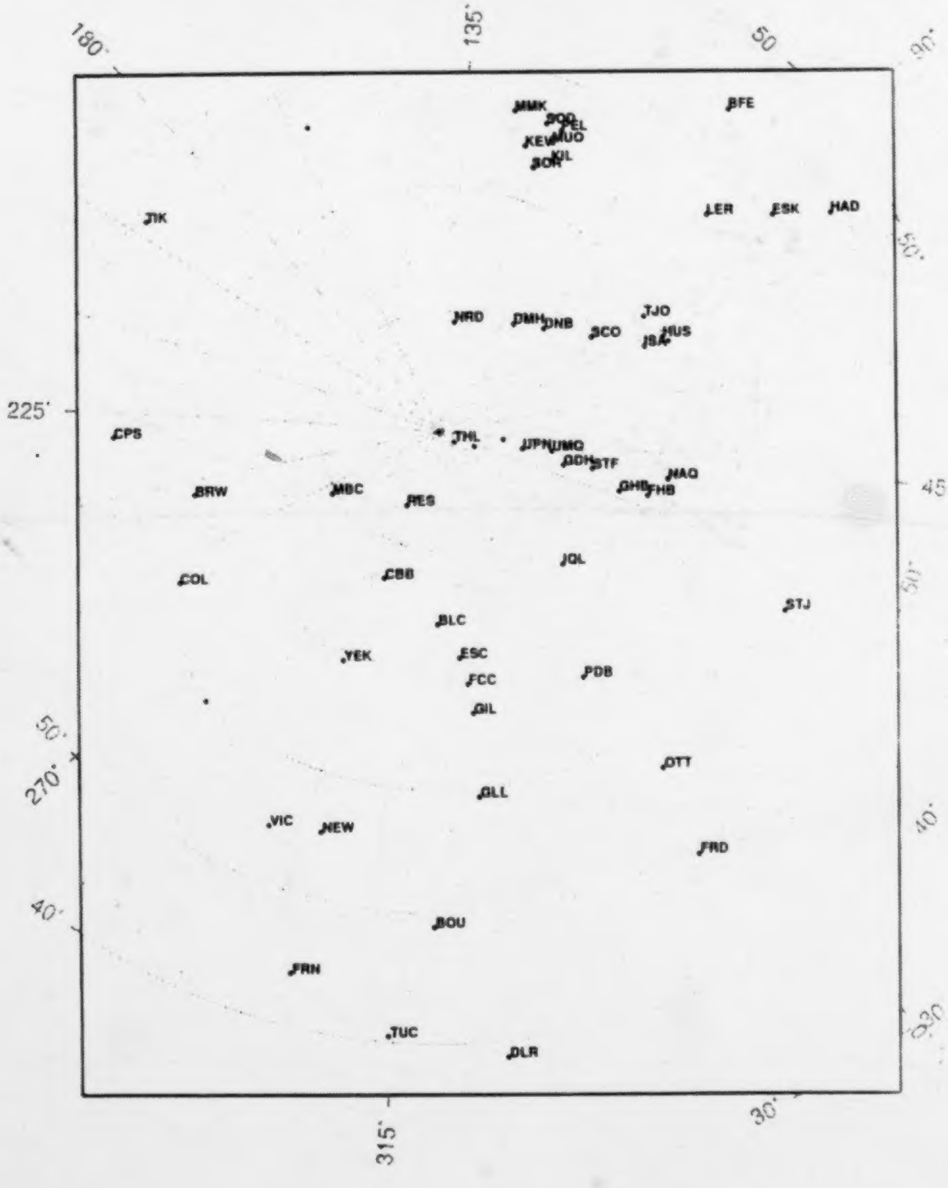
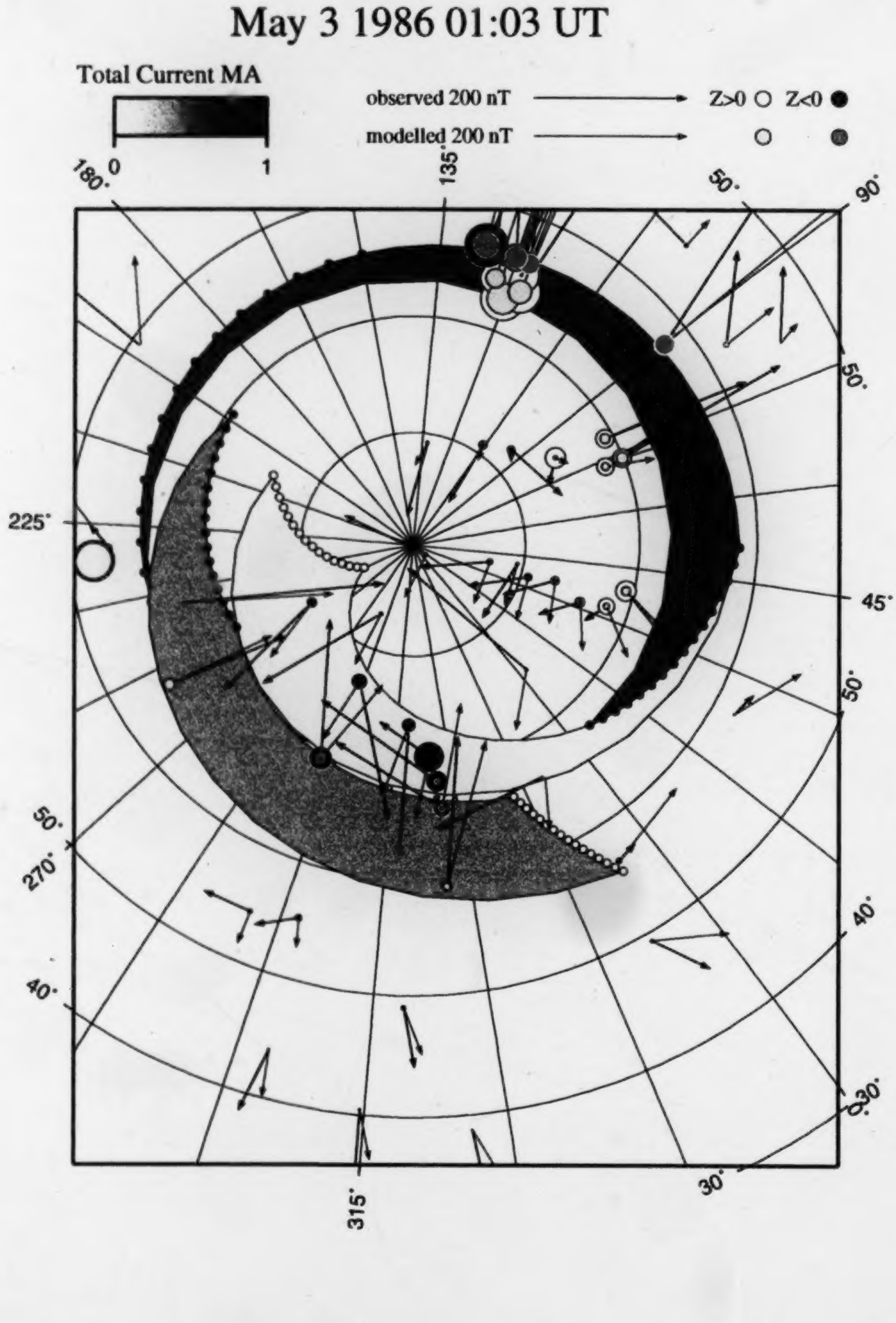

180°
135°
50
90°
MMK
SOD
PEL
KEV
MUO
KIL
SOR
BFE
TIK
LER
ESK
HAD
50°
HRD
DMH
DNB
SCO
TJO
ISA
HUS
225°
CPS
THL
UPN
UMQ
GDH
STF
NAQ
GHB
FHB
BRW
MBC
RES
45
JOL
COL
CBB
50°
STJ
BLC
YEK
ESC
FCC
PDB
GIL
50°
OTT
270°
GLL
VIC
NEW
40°
FRD
40°
BOU
FRN
TUC
DLR
30°
30°
315°

Figure 8.25 Global modelling results at 0103 UT May 3 1986.


May 3 1986 01:03 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●
180°
135°
90°
50°
50°
225°
45°
50°
50°
270°
40°
40°
30°
30°
315°

Perturbations at Ottawa and Saint John's were not accurately modelled by this system and appear to largely reflect the net upward current of the westward electrojet. The inaccuracy of the modelling of these stations will be dealt with in more detail below. Nevertheless, it appears clear from the modelling result that luminosity at the lower latitudes is associated with upward field aligned currents of the evening sector eastward electrojet, at least in the Ottawa region. One of the images in Figure 8.3 shows this luminous region, and as noted above, the luminosity is clearly at that location and not only apparently there due to foreshortening. Westward perturbations at Victoria and Newport may indicate field-aligned current flow strengthening the electrojet although there is no other evidence of this. Other midlatitude perturbations appear relatively well modelled: the Dst-like parameter in the model is -49 nT while the actual Dst parameter for this hour was -60 (see Figure 8.13). On the poorly constrained dayside, Tixie shows a negative perturbation which the model does not follow. This nearly 300 nT negative value indicates that the westward electrojet was at lower latitude on the dayside than the map indicates. It has been found for other events that there was not much evidence of the electrojet at a comparable position. For some other times in this run, the electrojet retreats to nearly Scandinavia, but this is clearly not correct considering the continuing presence of a negative perturbation at Tixie. The remedy for this would presumably be to reweight Tixie, but this would give one station too much dominance. It is also likely that the latitude of the electrojet changes so that a 'break', much as had to be introduced in Chapter 6, would have to be introduced. Overall, it seems more reasonable to simply be aware that for this event there was evidence of the westward electrojet near noon.

Figure 8.26 shows the situation very near the time of the 0111 onset. In general terms it is similar to that prevailing at 0103 UT. The changes in observed perturbations between the two times are small and the changes in the model are also. It would be difficult to attach much significance to the seeming poleward motion of the current system associated with the wedge. Certain unsuitable aspects of the model fit were thought to be susceptible to improvement should Region 1/2 field-aligned current systems be added. The poor match at Ottawa and St. John's has been mentioned in regard to the situation at 0103 UT and this match has not improved at 0112. In the manner previously discussed, north-south currents were added to the model and their magnitude varied while the other parameters (geometry and electrojet current) were held constant. The result of adding the north-south currents is shown in Figure 8.27.

Figure 8.26 Global modelling results at 0112 UT May 3 1986.

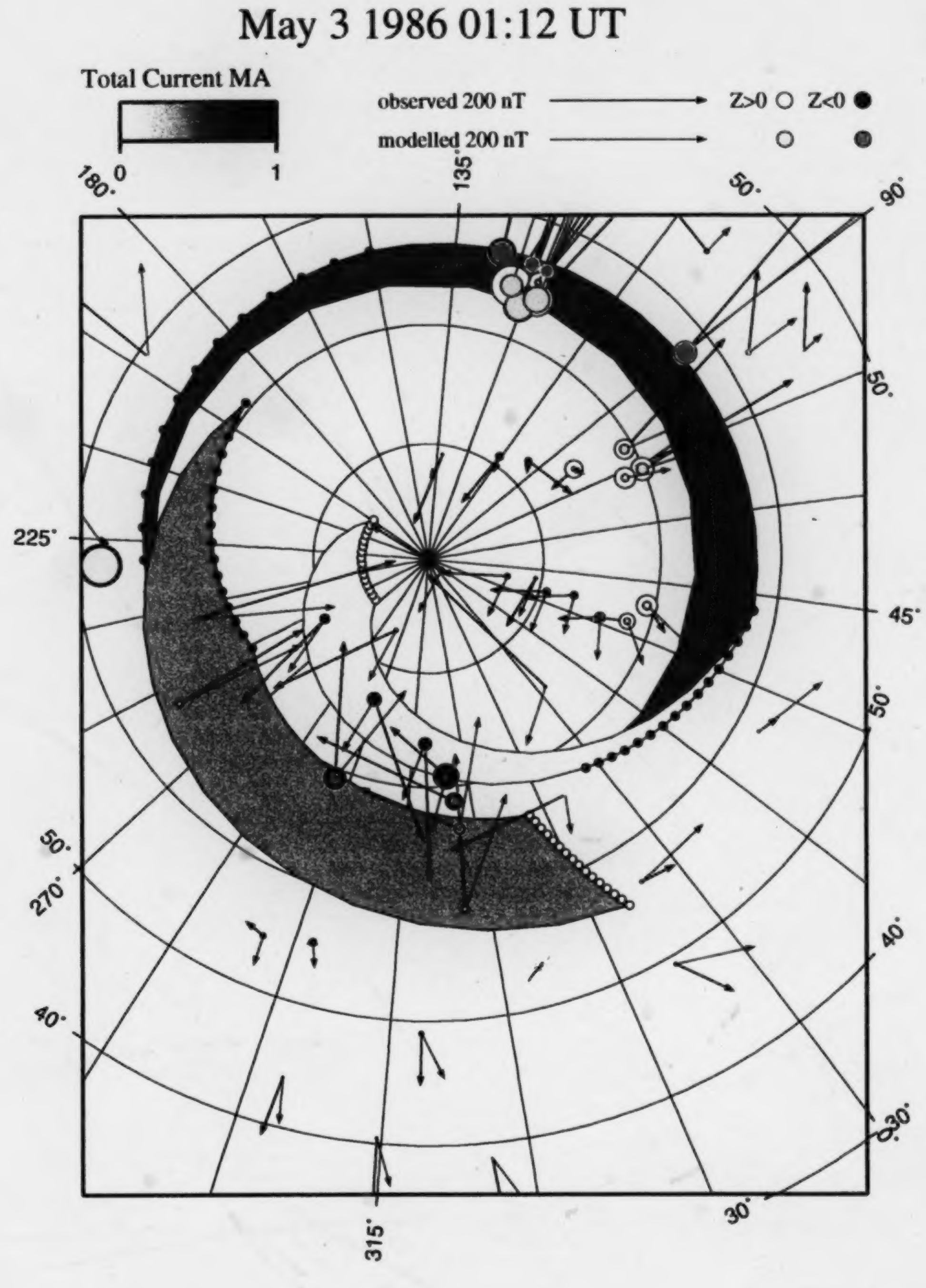
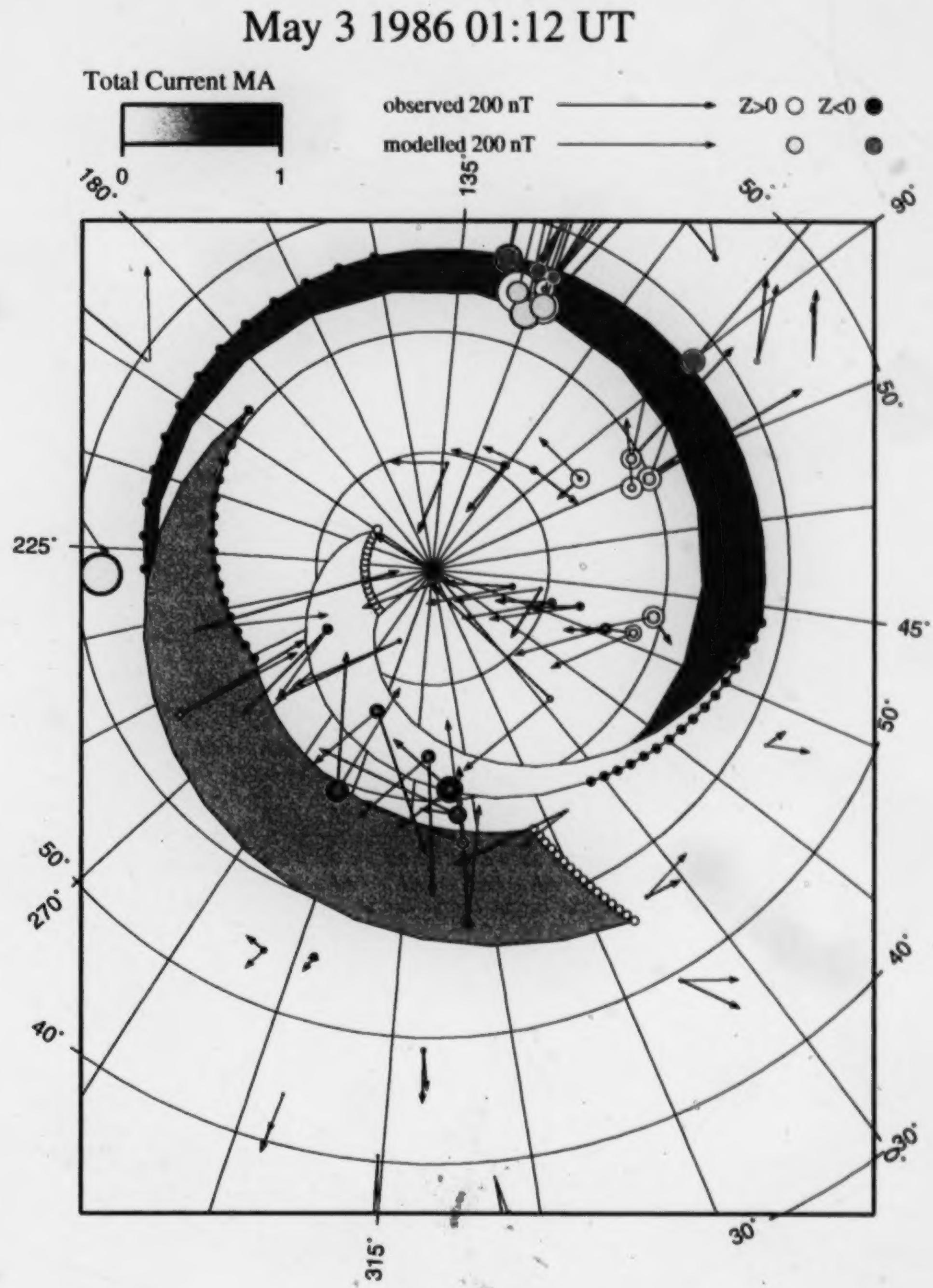

May 3 1986 01:12 UT
Total Current MA
0
1
observed 200 nT
Z>0 ○
Z<0 ●
modelled 200 nT
○
●
180°
135°
50°
90°
50°
225°
45°
50°
50°
270°
40°
40°
30°
30°
315°

Figure 8.27 Modelling results including NS currents at 0112 UT May 3 1986.

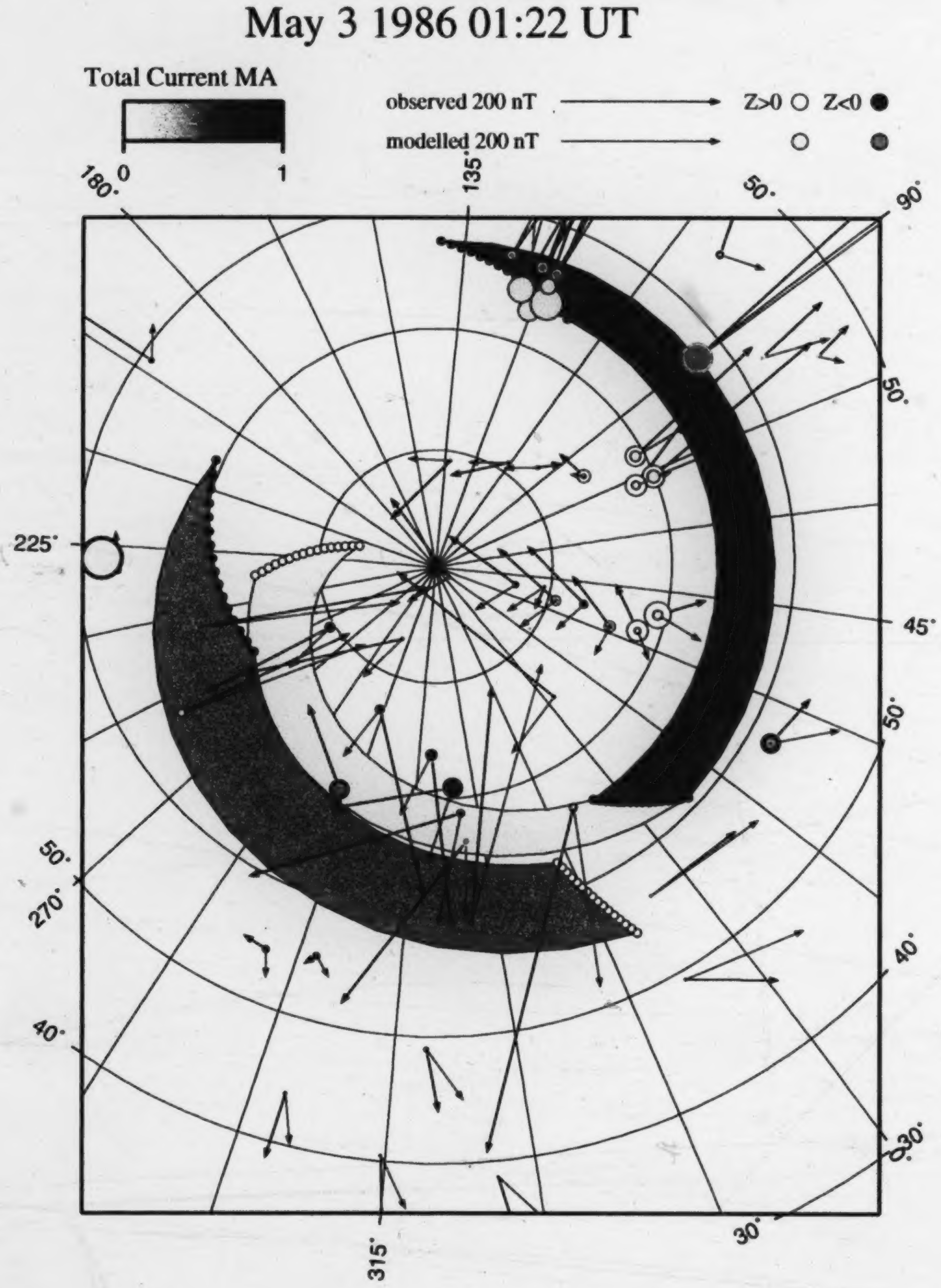

May 3 1986 01:12 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●
180°
135°
90°
50°
225°
45°
270°
40°
30°
315°

There is a notable improvement in the values at Ottawa and St. John's as a result of 'end effect' from the south-north Pedersen currents of the evening sector electrojet. These were found to total 0.81 MA in the south to north direction, and the region of the northward trending wedge had about 4.0 MA given to it in the same direction. All subauroral stations in North America had their match between data and model improved by the addition of these currents, although Victoria and Newport only marginally so. Other regions saw improvement in the data-model match as well. Improvements within the evening sector electrojet itself were not significant. Near its east end, the perturbation at Poste de la Baleine became much more westward, in better agreement with observation although the magnitude of the perturbation was not matched. The strong westward perturbations in the Churchill chain were not matched significantly better through this change, and little change in the Z component was seen there either. In the morning sector subauroral region, the eastward perturbations due to addition of north-south current flow (3.25 MA total) greatly improved the match at Hartland and Eskdalemuir, and the direction at Lerwick was improved to the point that the figure covers the observation arrow with the model arrow (although the magnitude of the latter is about 20% too small). At Brorfelde, the addition of a significant eastward component greatly improves the match to observation. Within the electrojet in the EISCAT region it appears that the veering eastward of the model arrows is slightly too great. The appropriate treatment of this area with regard to declination is disputable as indicated above, so not much weight is attached to this worsening of agreement. In the polar cap, results are dramatically improved at Resolute, and the whole of Greenland takes on a poleward (sunward) orientation although in general the magnitude of perturbations due to the field aligned currents is too large in western Greenland, and the direction not correct in eastern Greenland. As noted above, an evening sector orientation (south to north) was given, with a rather large current, in the northward trending wedge. This may be too large and too close to polar cap stations: it is likely that such a current would actually be associated with the evening sector electrojet. It is uncertain why the optimization in this sector produced a large current in the wedge region than in the electrojet. Nevertheless, the addition of the field-aligned currents appears to have improved the modelling and supports the idea that they contribute in a way observable in ground data. The linear current densities inferred from the above are 0.58 A/m associated with the northward 'wedge' system, and 0.25 A/m associated with the westward electrojet itself. The former is larger than that found in Chapter 6 from DMSP data during an active time, or than values quoted by Fujii *et al.* [1994]. However, this current has been shown to have produced overly large values of polar cap perturbations, and is almost certainly not located where the model placed it. A current of slightly smaller linear density, colocated with the evening sector electrojet, is the likely source of the polar cap perturbations. The current density associated with the westward electrojet is very similar to the canonical values. As was seen, its inclusion accounted very well for subauroral perturbations in Europe.

The post-onset observations and model results are shown in Figure 8.28. North-south currents are included in the model. It is clear that the region of evening sector westward current flow in the model is not well placed: stations with significant southward perturbations are north of it rather than nearly under it. Stations in this northerly region have reacted to the onset with some strengthening of their perturbations, but not a major strengthening. Thus, it seems inappropriate to refer to this northward trending region as the current wedge. However, there definitely is a region of westward current north of the evening sector eastward electrojet. That electrojet has strengthened somewhat, but again only by about 10%, compared to the pre-onset condition. The region of most interest is near Poste de la Baleine. This station, as has been discussed previously, shows the clearest signs of the 0111 onset, and the current wedge electrojet is near the station. The routine has not forced the intense current of the westward electrojet to actually overlie this station: likely it would were the station weight increased or were the location of the western part of the electrojet put in as an only variable during the run. Nevertheless, the current does not likely extend westward much past this station. There is strong evidence of an upward field-aligned current in the eastward perturbations at Ottawa and at Fredericksburg, which are very well modelled. The westward current is now indicated to be .75 MA in the electrojet, which is comparable to that found at this time in electrojet modelling using EISCAT stations. The most consistent picture is that the onset region was in the western part of the morning sector electrojet, and not in that section of westward current which is north of the evening sector electrojet. At this time, the picture of a monolithic westward electrojet, including the onset region, appears to be consistent with all observations. The question might then be asked whether the substorm current wedge existed in the sense of a region of downward field-aligned current, ionospheric current, and then upward field-aligned current. Certainly the latter two elements are there. There appears to be little evidence of a region of downward field-aligned current further east. This may be due to the lack of midlatitude stations (due to the presence of the ocean) to indicate it through perturbations to the eastward component. Of course, at this particular time, there is no need for such a field-aligned feed in any case since the electrojet current appears to be the same in both east and west. We thus may proceed to suggest that the 'wedge' in this case was primarily simply a prolongation of an existing westward electrojet. It should be noted that the general region of field-aligned current corresponds to regions of brightening in the images. Another somewhat unusual aspect of this event is that the onset-associated brightening moved eastward with time (this is not clear from the modelling nor from images shown above, but can easily be seen in rectified images in Baker *et al.* [1993]).

This completes the discussion of this event using global modelling. The technique has proven useful despite the fact that there are obvious flaws in the solutions.

Figure 8.28 Modelling results including NS currents at 0122 UT May 3 1986.


May 3 1986 01:22 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●
○ ●
180°
135°
50°
90°
50°
225°
45°
50°
50°
270°
40°
40°
30°
30°
315°

d. Summary

In this chapter there has been general discussion of data and presentation of new data relevant to the CDAW 9C 0111 onset which was discussed by Baker *et al.* [1993]. Use of meridian chain modelling was hampered by insufficient latitudinal coverage by the chains, especially to the south of Greenland. Global modelling was done and aided in understanding the layout of the current systems. Important results include:

1) a large onset took place at roughly 2315 on May 2, establishing a northern border current system in Greenland

2) the evening sector eastward electrojet was very active near the time of the 0111 onset. This partially accounts for luminosity seen in the Ottawa region which appeared to Baker *et al.* [1993] to be in conflict with the position of currents inferred from magnetograms

3) the currents flowing at the time of the 0111 onset were near Poste de la Baleine and were associated with an onset toward the southern border of the auroral oval near the midnight sector

4) there was an intensification or onset at about 0030; this onset may have been responsible for creation of a high latitude current system as the remains of an 'eye' of a substorm. These northerly currents were noted by Connors and Rostoker [1993] without explanation

5) the major onset at 0111 was followed by a second onset at about 0125

6) GOES 5 and GOES 6 likely detecting upward field-aligned current originating in the region between Poste de la Baleine and Churchill after onsets,

7) Ps 6 pulsations were present in the morning sector essentially throughout the active period and not only in 'recovery phase',

8) there was a northwest-trending westward electrojet located poleward of the evening sector eastward electrojet, which increased its current magnitude at the time of 0111 onset. However, the onset appears to have been localized in the electrojet in the midnight sector rather than primarily in this high-latitude evening sector electrojet, and

9) strong westward perturbations in the dusk sector auroral oval near onset time.

From a technical standpoint, several conclusions about modelling could be drawn:

1) with less than half the electrojet over a meridian line of stations, it is difficult to obtain reliable results, even if there is constraint from a faraway station

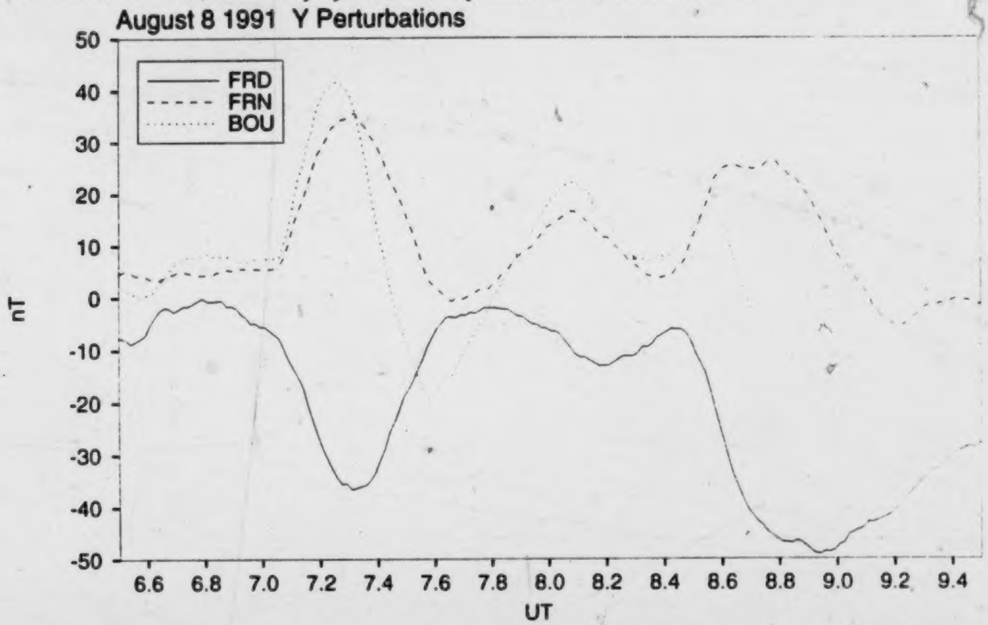
2) Ps 6 pulsations did not notably interfere with large-scale modelling

3) that addition of north-south current systems (including both ionospheric and field aligned current flow) of reasonable magnitude substantially improved the agreement between models and observations at mid-latitude stations and in the polar cap, without much affecting the agreement in the auroral oval. Latitudinal placement of such currents by the routine did not appear accurate in that it located some of them in the high-latitude evening sector westward electrojet when they would more plausibly be located in the evening sector eastward electrojet.

Chapter 9. August 8, 1991

This event was chosen after examination of USGS mid-latitude one-minute magnetic data. A pronounced and symmetric transition bay was seen in central North America at roughly 0700 UT (Figure 9.1) on August 8, 1991. Availability of CANOPUS data to supplement USGS data gives good coverage in North America, and STEP and INTERMAGNET CD-ROM data fill this in further and give the possibility of a global view. MARIA meridian scanning photometer data in several spectral bands give the potential of studying the connection of luminosity and magnetic signatures.

Figure 9.1 Midlatitude Y components on August 8, 1991. North American stations are shown, with very symmetric bay from 7.0 to about 7.6 UT.



Initial modelling of this event was done using a simplified fitting routine in Cartesian geometry [Connors, 1993]. This routine is the predecessor to the one used in the present study. Using linear current segments, the Levenberg-Marquardt algorithm was implemented using analytic expressions for both the field and its derivatives. The intention was to evaluate the algorithm in a magnetic modelling application while retaining analytical expressions amenable to debugging. The substorm current wedge was characterized by the electrical current positional parameters giving the y-coordinate of the ionospheric path and the x-coordinates of the downward and upward field-aligned currents, which were given an estimated inclination and produced ground effect at mid-latitudes. Stations in the USGS network were chosen having approximately the same (40° N) magnetic latitude, considered to be the x axis. The positional parameters and the current required to produce a good fit of the mid-latitude data were determined. The model is similar to that used by Horning *et al.* [1974] as a basis for modelling using linear inverse theory. In practice the methods functioned well and it was felt well advised to proceed with full three-dimensional modelling as described in this study.

a. Data Preparation

As USGS one-minute CD-ROM data were duplicated on the INTERMAGNET CD-ROM, all stations on that disk were read into files for the day of August 8, 1991. Similarly, all data from the STEP CD-ROM were read. CANOPUS data were taken from tapes and converted to one-minute values on a VAX, then transferred to the computational machines, where they were plotted and compared to printouts. By examination of USGS data on a PC, quiet days of August 13 and 14 were found. Initially, August 13 seemed suitable, but plotting of smoothed data from this day showed some bias toward negative X values at midlatitudes. It was confirmed that August 12 was moderately disturbed and that D_{st} was (likely due to persistence of ring current) somewhat negative in the earlier part of August 13, so August 14 was used as a baseline. This choice was later confirmed by examination of the K_p activity index for these days. Files for August 14 were prepared in much the same manner as for the day under study. These were then smoothed by box-averaging with a one hour window, and resulting XYZ values subtracted from the study day data values at each station. There was, in fact, activity in the Churchill sector of CANOPUS on August 14 (ca. 100 nT onset-like activity) and this was eliminated by using a six-hour smoothing on affected comparison data. Since data were not available for August 14 for the CANOPUS station Island Lake (ISLL), it was determined that perturbations at 6 UT on the study day were close to zero at other stations, and perturbations at this station were determined by subtracting the one-hour averaged value at 6 UT from the other values recorded that day. Z component data were not available at Dawson (DAWS), for either August 13 or 14, so XY values were prepared as for the other CANOPUS stations with good data, but Z was referenced to 6 UT as for ISLL. Consistency checks were made of the data for duplicated stations from various sources, and the data being found to be consistent (although very small differences did exist, likely due to rounding for data storage), duplicate stations were removed from the list of files to be used in modelling. The CANOPUS station at Fort Churchill did not have data available but this station is very close to (200 m away from) the Geological Survey of Canada station and those data were available on the CD-ROMs.

With the noted exceptions, CANOPUS data in general were of very high quality. A data dropout of roughly two minutes duration was present for unknown reasons in the one-minute data produced (on a VAX in 1993) for Pinawa, the southernmost station of the Churchill meridian line. This took place almost exactly at onset time. For modelling purposes, it was felt justifiable to remove the gap by linear interpolation from adjoining data. Inspection of the data from the adjoining stations Island Lake and Glenlea, very close to Pinawa, confirmed that linear interpolation was a reasonable choice. Plots of the original 5-second CANOPUS data confirmed this.

Initial modelling results indicated that the evening sector electrojet could be of some importance to the study, so more data were sought for that sector. The STEP CD-ROM had Russian data only from Tixie, and efforts to obtain other Russian data were not fruitful. It was possible to obtain University of Alaska perturbation data, which unfortunately had numerous dropouts, from the U. S. National Geophysical Data Center[1] web page. Data from the 8 to 14

---

[1] URL is http://www.ngdc.noaa.gov:8080/production/html/GEOMAG/geo_search_frames.html

August were obtained so that quiet day removal could be done as for the STEP and CANOPUS data. A complication arose in that these Alaska data were in variation form but stored as D, H, Z. With D in units of 0.1 arcminutes from an unknown reference declination, and horizontal and vertical intensities H and Z given only as nT offsets from an unknown level, this information is not sufficient to produce the XYZ (ultimately $X_M Y_M Z$) perturbations needed by the routine used here. To bypass this problem, an attempt was made to find measured absolute fields at the Alaska stations. When these were not available in standard sources for all of them, or dated from about 1957, it was decided to simply add in IGRF90 values calculated from a computer program [Rod Farlee, 1993, personal communication of computer program based on Langel, 1987]. By adding IGRF values to the declination perturbation and to that of the horizontal intensity, a pseudo-absolute reading was created, and from this the X and Y components could be calculated in the usual manner. The pseudo-absolute Z component was obtained by directly adding the IGRF value to the perturbation. It was verified for several stations with known XYZ values that the computed IGRF values for XYZ were close to those on the quiet day. It was further verified that differences in perturbations at the stations AMU (Anchorage) and TLK (Talkeetna), which are geographically very close, were small. This was true throughout the UT day, and at 7 UT the X and Y component differences were in fact less than 5 nT, while the Z difference was about 80 nT. The latter is not considered significant due to the latitude difference of the stations and the sensitivity of the Z component to latitude near the electrojet, as discussed in Chapter 3. The implicit assumption that the calibration of the Alaskan stations did not change between the day studied and the reference day of 14 August seemed substantiated by the reasonable appearance of the perturbation files obtained after smoothing the reference day and subtracting, much as detailed above for other stations. For reference purposes, details of the stations treated in this manner are given in Table 9.1.

Table 9.1 Parameters of Alaskan Stations

	Name	Lat	Lon	Dec	H nT	Z nT	X nT	Y nT
AMU	Anchor-age	61.2	210.1	24.1	15296	53442	13964	6244
AVI	Arctic Village	68.1	214.4	30.2	10452	56468	9034	5257
FYU	Fort Yukon	66.6	214.7	29.5	11362	56149	9905	5566
INK	Inuvik	68.2	226.7	36.1	8593	57854	6947	5057
SAH	Sachs Harbour	72.0	235.0	44.0	5168	58438	3720	3589
TLK	Talkeet-na	62.3	209.13	24.5	14657	53916	13337	6080

Digital one-minute data for Glenlea and some other stations, including many from the southern hemisphere, were also obtained from the NGDC web source. Data from Glenlea were in absolute form and prepared much as described above for CANOPUS data. Most of the other stations were conjugate stations and used for reference purposes only.

Finally, after initial wedge modelling as detailed below, it was judged useful to obtain morning sector data from Greenland. This was sought largely since indications were that the morning sector perturbations were rather weak: further, the availability of chain data would allow good electrojet modelling and a gap would be filled for global modelling. These data were obtained by ftp from the Danish Meteorological Institute with 20 second spacing, for the study day and for the comparison day of August 14. An existing program was modified to read the 20 second data and decimate it to one minute. The program output was compared to ASCII values in the original files to ensure that proper decoding had occurred.

b. Chronological Summary

The IMF turned southward in the late UT hours of August 7, but IMP-8 IMF values on August 8 are not available after about 3 UT and thus are not shown here. After steady southward field for many hours, there was major substorm activity with onsets during the period 1 to 4 UT, as evidenced by negative perturbations at Narssarssuaq (NAQ) and at the highest latitude CANOPUS stations. Brief onsets took place between 70° and 75° while these stations were in the evening sector. This is reminiscent of activity noted well prior to major onsets during the CDAW 9B event discussed in Chapter 6. Unlike in that event, there appears to have been evidence of the eastward electrojet at lower latitude stations in the evening sector. This activity declined gradually until about 0530 UT, when there is again evidence of activity at high latitude stations. This includes changes of up to 200 nT at high-latitude Greenland stations in the morning sector and onset-like intensifications at Antarctic stations also in the morning sector. It should be noted that each magnetogram acquired showed further evidence of activity which had not been evident from the North American data. Details of the morning sector activity are not further discussed here except to comment that it was at high latitudes and coincided with what had been interpreted to be the beginning of a new growth phase. The high latitude evening sector showed an intensification of the eastward electrojet at Barrow at 0530 UT and an abrupt decline in the Z component there at 0600 UT which may signal its southward motion. At approximately 7 UT, evidence of the major substorm onset at relatively low latitudes, which is the focus of this modelling study, was seen in the magnetic data. Table 9.2 summarizes the progress of the event. Absence of IMF data does not allow comments to be made about triggering.

Inspection of data from high latitude evening sector stations AVI and FYU confirmed the presence of the eastward electrojet there after 0530 UT and its southward motion until onset time. Its maximum perturbation was about 200 nT in the X_M component. At 0530 UT negative perturbations intensified at station NAQ at comparable magnetic latitude in southern Greenland (morning sector), reaching about -200 nT by onset time. The growth phase thus appears to have resulted in the formation of relatively symmetric driven system electrojets.

Table 9.2 7 UT Onset on August 8 1991

Time (UT)	Activity	Observational Basis
0-3	IMF has been southward for hours	IMP-8
1-4	onsets in midnight sector (Atlantic) and evening sector	NAQ, high-latitude CANOPUS
0530	growth phase begins	NAQ, GOES-5, GOES-6
0630	evening sector electrojet growth and southward motion	CMO, BRW magnetograms
0700	optical onset (65° Invariant)	Pinawa, Gillam MPA
0703	magnetic onset	FRN, BOU, modelling
0705-0710	dipolarization at synchronous orbit	GOES-6, GOES-7
0703	Pi1 and Pi2 onsets	Pinawa, Gillam magnetometers
0740	end, midlatitude wedge signatures	(transition bay at BOU ends)
0830	second onset	magnetometers, MPA, GOES

GOES-6 (see Figure 9.2) was at 135° W geographic longitude (footprint near station DAWS) and observed the growth phase mainly through a decrease in the H (northward) component prior to onset. Already depressed from the quiet day level, this component clearly began to decline once more at 0530 as the electrojets strengthened. This behaviour is consistent with ground observations near the footprint point. At DAWS, the data are not suitable for use in modelling due to dropouts and uncertainty in baseline values. However, the period from about 0520 to 0600 clearly shows an increase of about 150 nT in the X component. A similar but not quite as large increase takes place at the nearby station College (CMO). Figure 9.3 shows data from GOES-7 at 98° W, somewhat west of the CANOPUS Churchill line. At this spacecraft, the growth phase was marked by strengthening of the radially outward component with little change in H. Both growth phase signatures indicate strengthening of cross-tail current, although they are more typically observed together on one spacecraft. GOES-7 shows the typical onset signatures of H component increase and radial component reduction (i.e. dipolarization) at onset time, but the dipolarization appears to be short-lived. At GOES-6, further west, the dipolarization signature in the northward component is not accompanied by as clear a signature in the earthward component. The most definite dipolarization signatures at GOES-7 take place at 7.05 UT although there is some change in the northward component already at 7.00 UT or earlier. At GOES-6, there appears to be a small dipolarization at 6.8 UT. The main dipolarization is abrupt, with radial component changes starting at 7.05 UT, but the northward component increase mainly taking place at 7.10 UT.

Figure 9.2 Perturbations at GOES-6 on August 8, 1991. Components are, from top, northward, earthward, and eastward. Smooth lines are comparison traces from August 13, box-smoothed with a six hour window.

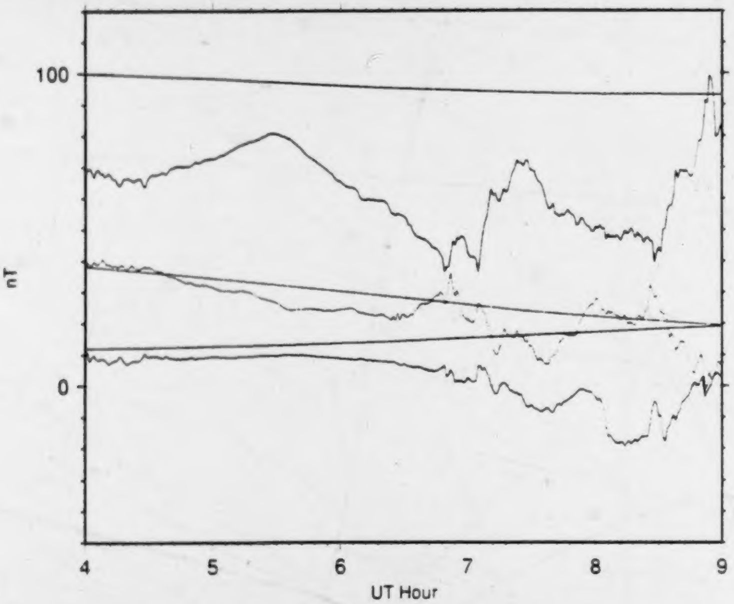
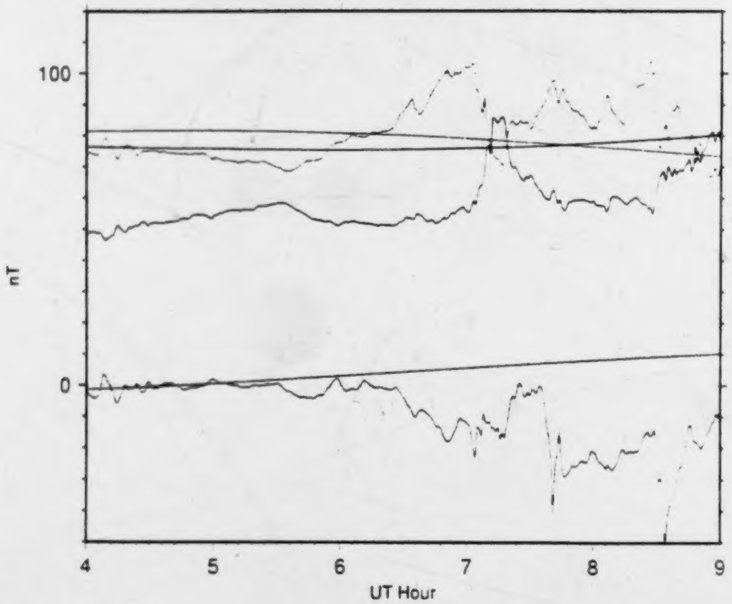


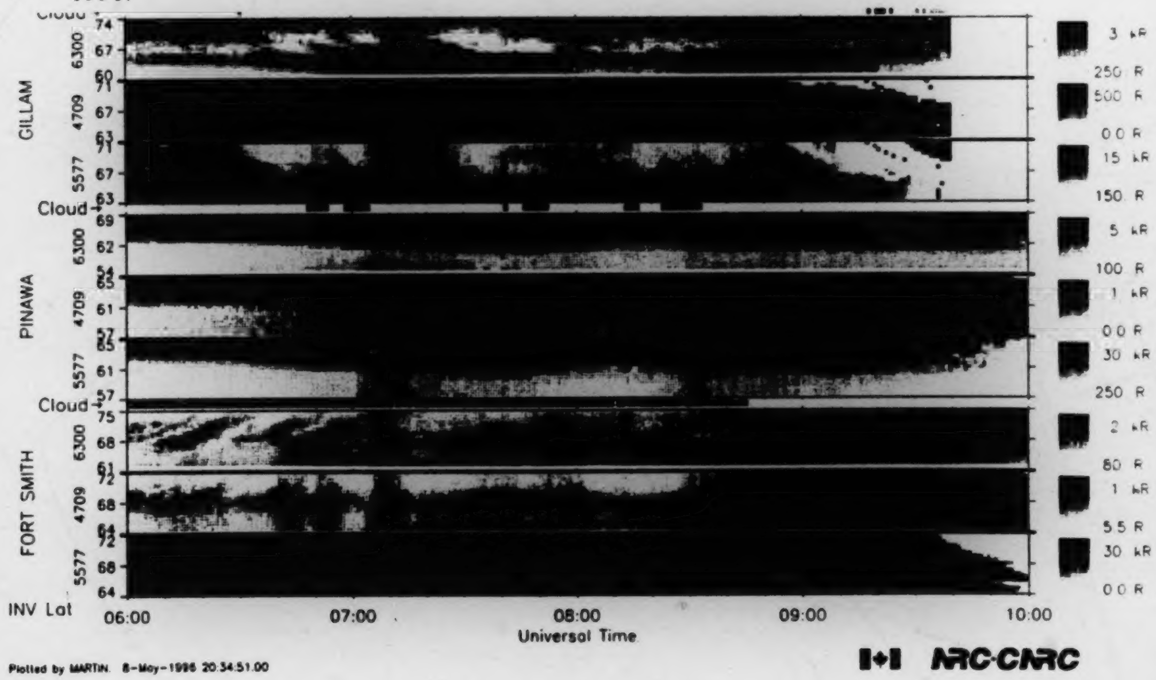
Figure 9.3 Perturbations at GOES-7 on August 8, 1991. Components are, from top, earthward, northward, and eastward. Accompanying smooth lines are comparison traces.



Later activity at the GOES satellites includes more complete dipolarizations, but it is only as the spacecraft approach the dayside, after about 14 UT, that the observed perturbations become negligibly different from those of the comparison day. At both spacecraft, the eastward component, as compared to the quiet day, is negative at all times after the start of the growth phase. This is the direction to be expected in association with westward electrojets although, as noted above, GOES-6 has its footprint in a region of eastward electrojet until onset. Prior to onset, it is likely that the westward field at GOES-6 is at least partly due to the more tail-like nature of the field at synchronous orbit.

The midlatitude perturbations used as a selection criterion for this event may be inspected in more detail. Figure 9.1 shows the Y component perturbations at the midlatitude stations initially used by Connors [1993]. Comparison of the Fresno and Fredericksburg traces shows that the wedge is well centred over North America, with positive Y at the former indicating upward field-aligned current north of it, and the rather symmetric negative Y perturbation at the latter indicating downward current. The curve at Boulder is a transition bay, with initial indications of upward field-aligned current being replaced by downward signatures toward the end of the event. As was clear from the GOES data, a second clear onset took place at about 0830, but it is not discussed in detail here. There may in fact be another distinct onset somewhat before 0800, also not within the scope of this study.

Figure 9.4 Meridian scanning photometry at CANOPUS stations on August 8 1991.



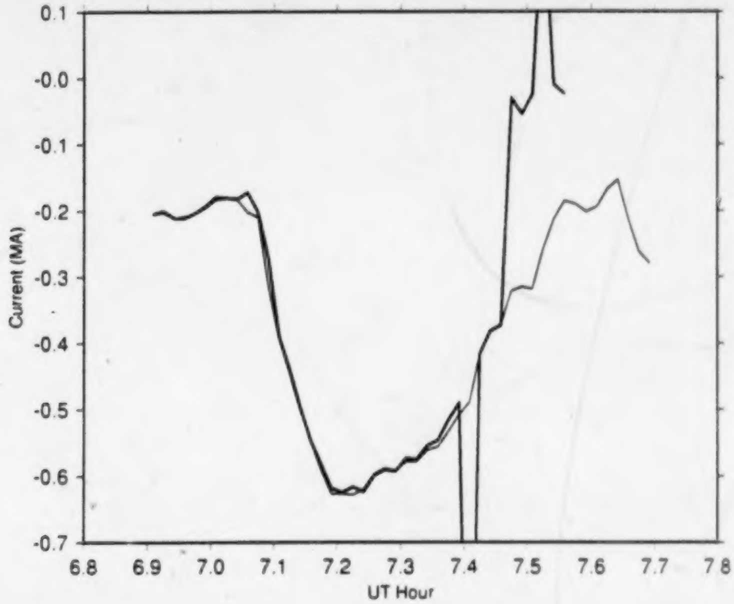
The basic character of the event may also be deduced by examining Figure 9.4, which shows meridian scanning photometer data from several CANOPUS stations. This type of data was briefly discussed in Section 3.c.4. Gillam and Pinawa, in the Churchill meridian, show the onset clearly just after 0700 UT. Both stations concur in placing the latitudes of activity, as indicated by emissions, between 63° and 67° in the PACE Invariant system. This is a useful check on modelling results. From Gillam, there is evidence of earlier activity at about 70° at roughly 0630, and in addition, there are equatorward moving features seen just before the 0700 onset. There may be some indication of the 0630 activity in the magnetic perturbations of Figure 9.1, but this activity will not be discussed further here. The onset just after 0700 has a clear counterpart in the midlatitude magnetic bays, as does the one at about 0830. The latter appears strongest at the most westerly station of Fort Smith in the optical data, whereas the 0700 onset did not produce much luminosity there.

c. Meridian chain results

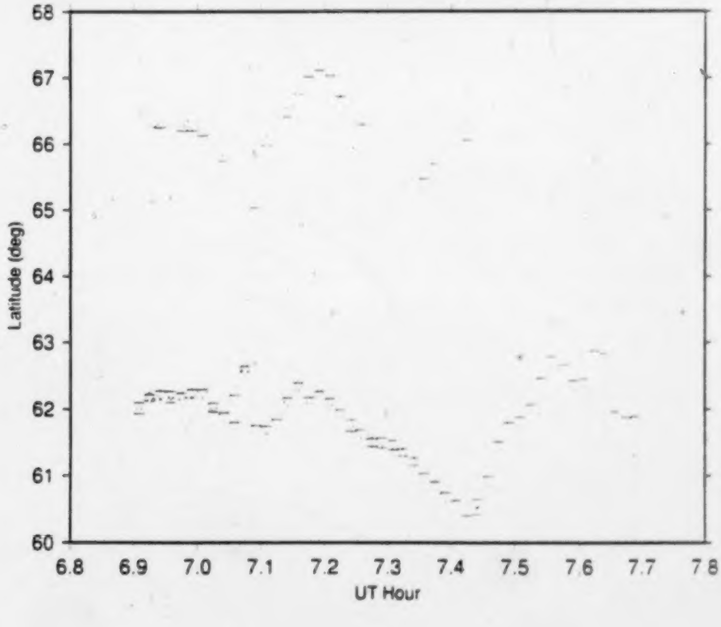
Meridian chain data were available for the study of the onset region near the CANOPUS Churchill line, and were later obtained to allow study of the eastward electrojet in Alaska. Data preparation was described above, and for the Churchill line, with near-zero declinations, the meridian chain modelling was done with unrotated data (XYZ perturbations).

A first attempt at modelling was done using CANOPUS data before final baselines were established. Since features of this first attempt illustrate the robustness of the routine, it will be discussed briefly. To simulate a long electrojet passing over the Churchill line (which is at roughly 325° or -35° centred dipole geomagnetic longitude), field-aligned currents fixed in longitude at -60° and 0° were joined by an ionospheric current whose magnitude, latitude and width were allowed to vary. By a fortuitous accident, the current in the initial condition for the first modelling run was reversed, an initial value of 0.5 MA eastward being specified. Convergence was good in any case, with the routine reversing the direction of current flow. Initial latitudes for the northern and southern borders of the electrojet were specified as 64 and 65 degrees. Generally the electrojet was found to be much wider than this. From this flawed run, a second run was done starting from the parameters of one of the best fit times. The results were nearly identical for both current and electrojet borders as shown in Figures 9.5 and 9.6, which combine the results from each of these modelling runs. The robustness of the Levenberg-Marquardt routine as used for meridian chains is demonstrated by these near-identical results from greatly different initial conditions. Later modelling proceeded with proper baselines and using the best fit from these preliminary runs as the initial condition.

9.5 Current over Churchill chain from 6.8 to 7.8 UT August 8 1991. Heavy line shows results from bad initial condition, thinner line from realistic one.

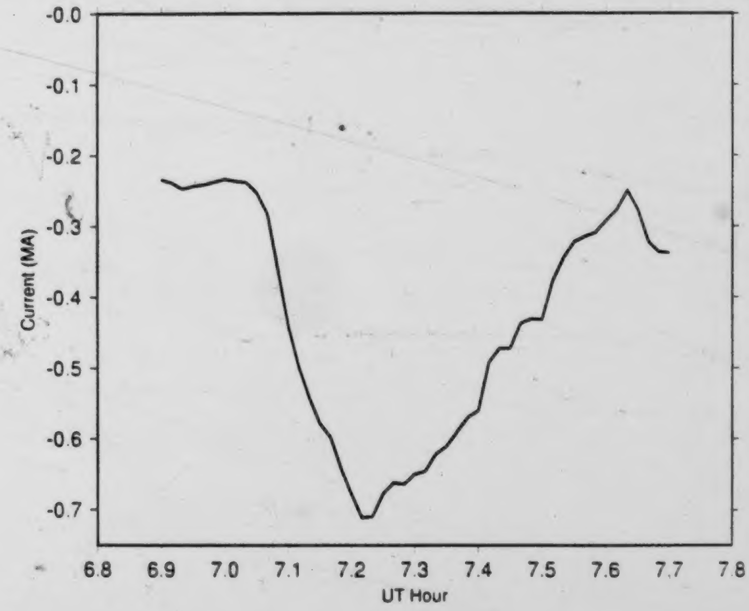


9.6 Borders of electrojet over Churchill chain, 6.8 to 7.8 UT August 8 1991. Upward triangles show result from poor initial condition, prior to 7.5 UT. Inverted triangles show result from improved initial condition (also run longer).



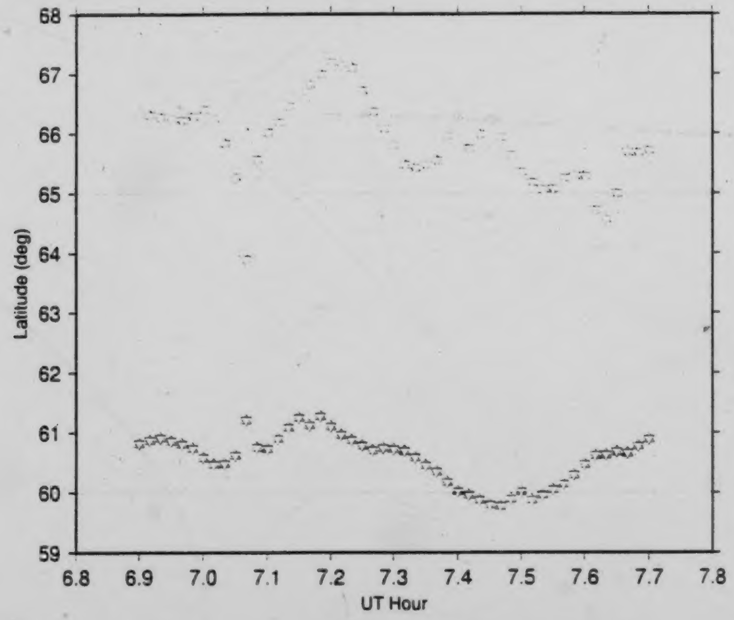
The most noteworthy result from these modelling runs is the narrowing of the electrojet near onset time. This effect was also noted in Chapter 7 using data from the Russian chain, and the claim was justified through the presentation of a series of observed/modelled latitude profiles. Here only two such profiles will be presented in the interest of brevity as justification of the claim of electrojet narrowing is detailed in the next paragraph. By time 7.1 UT the electrojet current has intensified markedly, having commenced its decline to more negative values at 7.05 UT in the best model. It is at 7.08 UT that maximal narrowing of the electrojet takes place, in other words, shortly after the time of onset and at the time that the current is declining most quickly. The 'necking' is very brief, lasting only about two minutes, although narrowing may have commenced as early as 7.00 UT, shortly before onset. Since the current has not yet changed when the necking starts, the effect seems to be associated with the start of onset. The narrowing decreases the electrojet width to roughly one degree, and is in the southern part of the electrojet. The subsequent strengthening of the electrojet is accompanied by poleward motion of the northern border, in the fashion of the 'eye' of the substorm and as detailed in Chapter 7 with images. The poleward extension in this case is small, at most two degrees above the preexisting electrojet border.

Figure 9.7 Current over Churchill chain from 6.8 to 7.8 UT August 8 1991. Improved modelling includes station GLL at southern end and has improved baseline.



The suggestion from the preliminary modelling that an interesting reconfiguration of the electrojet took place at onset time was verified by modelling with better baselines. In addition, station GLL (Glenlea), south of the southernmost CANOPUS station at Pinawa, was added to the data set. We will now examine in more detail the magnetic events at time of onset and what the fitting routine is telling us about those events. First, examination of Figures 9.7 and 9.8 will show that the results of this modelling with improved baseline and addition of a station are similar in all essential aspects to those arising from the poorer data using only CANOPUS stations. The strengthening of the westward current begins at 7.05 UT and its rate of strengthening increases as seen by the steepening of the curve showing (eastward) current over the meridian. Again the simple forward model suggests a narrowing of the electrojet at onset time. The effect is less pronounced than in the earlier models, but still results in a near-halving of the electrojet width. It is also confirmed to start at 7.00 UT, that is, 3 minutes before the strengthening of the current commences. The necking down of the electrojet between times 7.5 and 7.6 UT seen in Figure 9.6 is no longer present and can be attributed to poor baselining at a time when the current was weak, making the effects of baseline errors more important. The overall current as shown in Figure 9.7 is slightly greater in magnitude than that shown in Figure 9.5, a result partly attributable to baseline improvement since the baseline originally used was slightly depressed in the earlier modelling, resulting in lessened perturbations and thus less current. It is also likely that the very southerly location of the electrojet results in some current not being 'seen' when the southernmost station, Glenlea, is not included. In any case, there is substantial agreement between the best profile from Figure 9.5 and that of Figure 9.7.

Figure 9.8 Borders of electrojet over Churchill chain, 6.8 to 7.8 UT August 8 1991. Improved modelling with station GLL at southern end and improved baseline.



The modelled electrojet borders may be compared to the meridian scanning photometer data of Figure 9.4. Prior to onset there appear to be two arc systems as imaged at Gillam, which are moving southward at the same rate between about 0650 UT and onset time (and possibly afterward). At the start of modelling at 6.9 UT (0654 UT) the northerly arc is at 67° in invariant latitude, which is comparable to the 66° centred dipole latitude found for the northern electrojet border from the current modelling. The southern border of luminosity as imaged from Pinawa is at about 62° invariant, and 61° in the magnetic modelling, again consistent.

After onset, the northern border of luminosity moves northward to almost 69° invariant and the electrojet to over 67° centred dipole, this maximum occurring at 7.2 UT in both. This border moves steadily southward until 7.5 UT when activity appears at the northern border, and then remains fairly steadily near 65°. A brightening in the northern border (which has moved southward to about the original onset latitude) shows as a brief southward motion of the northern electrojet border at about 7.65 UT. There is excellent agreement between motions of regions of luminosity and electrojet boundaries.

The correlation of timing of auroral brightenings with the onset time is not so good as the spatial correlation between regions of brightness and electrojet borders. As detailed above, the onset time from magnetic data has been accurately determined as 7.05 UT (0703) from the magnetic modelling. Auroral brightening appears to have begun near the southern border of luminosity very shortly after 7.00 UT, at latest 0701 UT, as imaged both at Gillam and at Pinawa. This applies to auroral emissions at wavelengths 486.1 and 557.7 nm, and to the 625.0 nm background. Auroral emissions at 470.9 and 630.0 nm appear to come to maximum brightness slightly later.

Another approach to timing is through the use of pulsations. Both Pi2 (5 to 15 mHz band) and Pi1 (15 to 50 mHz band) pulsations are closely associated with substorm onset and with the initial field-aligned currents of the expansive phase [Samson, 1991]. Five second data were transformed to the Fourier domain and had Butterworth bandpass filters applied. The filter output was then transformed back to the time domain. Filtering for pulsations at Pinawa and Island Lake (Figures 9.9 and 9.10) showed a first half cycle of Pi2 already near its extremum by onset time at 0703 UT (negative half cycle at Pinawa, positive at Island Lake). Filtering in the Fourier domain is not causal and this apparent advance of the start of Pi2's does not necessarily indicate that the onset had taken place earlier than 0703. It is interesting to note, however, that the decline in X and Z components observed at Pinawa was already underway at 0700. Additionally, there is a positive excursion in the Pi2 band as measured at Pinawa, which peaks immediately after 0700. This may be associated with the initial auroral brightening. At both Pinawa and Island Lake, Pi1 pulsations started at 0703 UT and had reached full amplitude of about 20 nT by 0704 UT. In distinction to the Pi2 pulsations, the Pi1's at both stations were roughly in phase. Their amplitude declined significantly after 7.2 UT, which was the time of maximum current magnitude and maximum northerly expansion of the electrojet.

The foregoing suggests that there is a good correspondence between the borders of auroral luminosity and the borders of a simple electrojet model, a result which is not unexpected. The value of this result lies more in confirming the validity of results from the electrojet model than in being an original observation. The use of the electrojet model to specify onset time does lead to the useful conclusion that auroral brightening took place shortly before the onset of current flow or of Pi 1 band pulsations, and possibly was coincident with the start of Pi 2 pulsations.

Figure 9.9 Pi2 band pulsations at Pinawa and Island Lake on August 8 1991. Time units are UT in twelfth-minute units, so that the interval 4800 to 5520 shown corresponds to 0640 to 0740 UT. Major annotations are at 10 minute intervals.

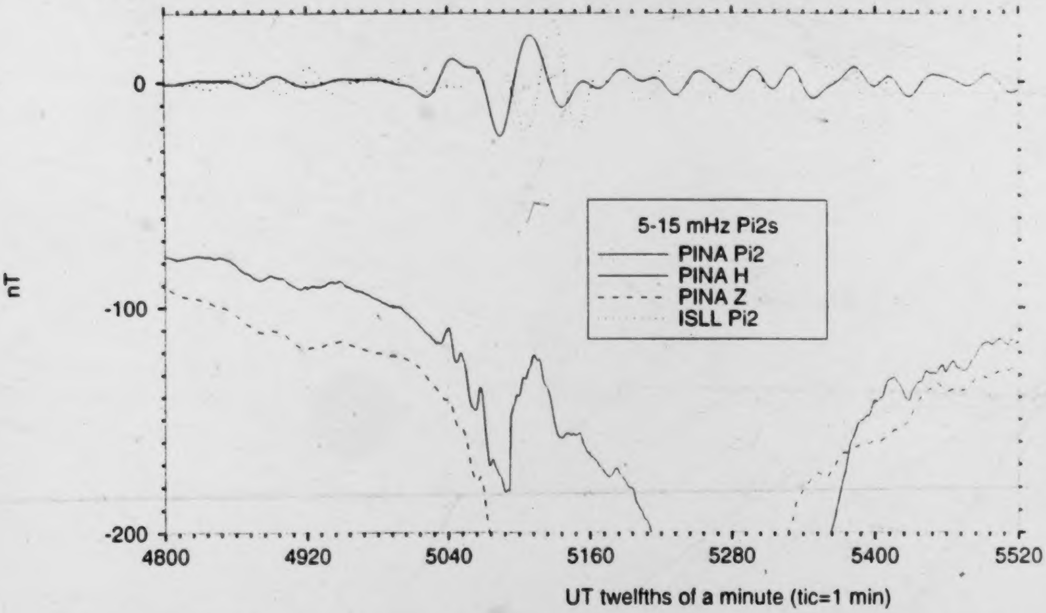
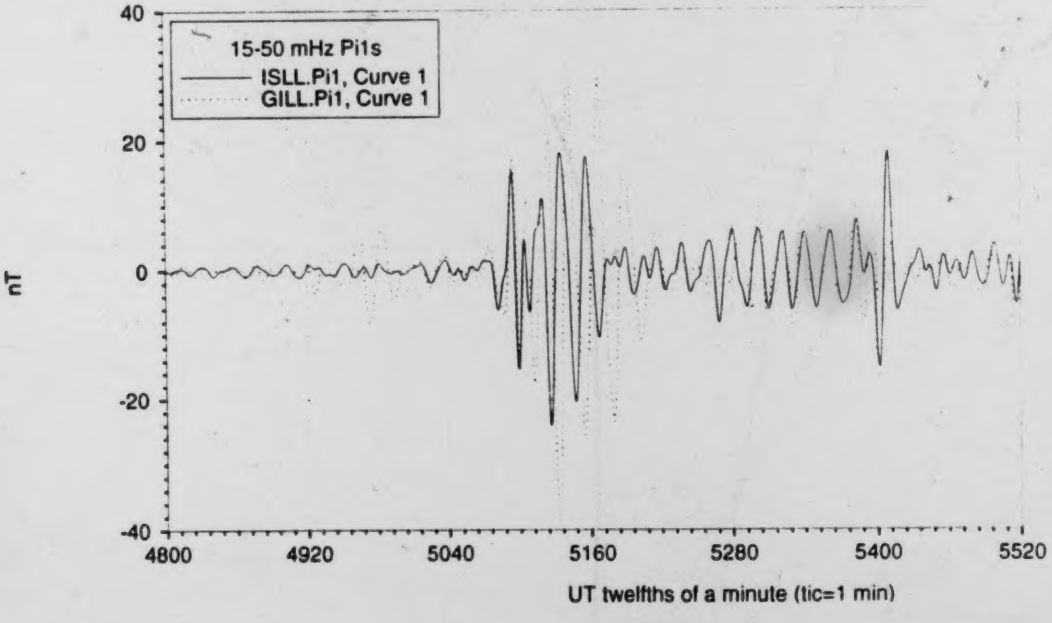
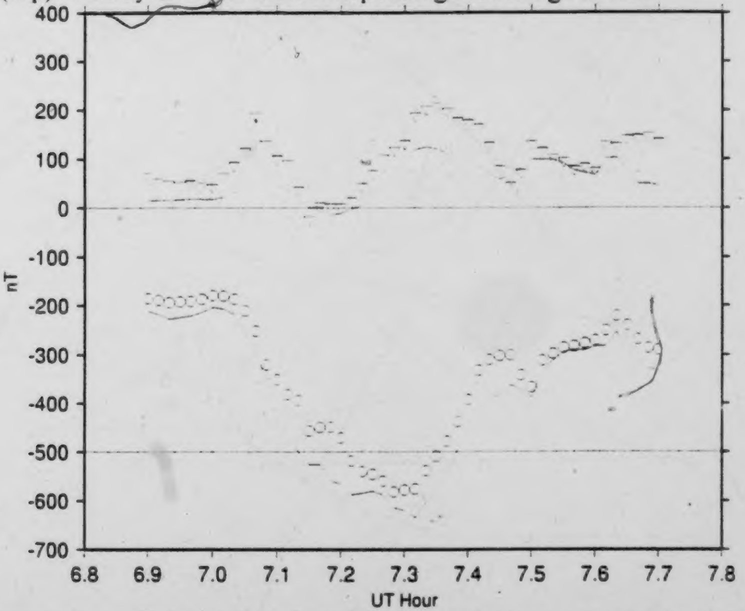


Figure 9.10 Pi1 band pulsations at Pinawa and Island Lake on August 8 1991. Time units are as for previous figure. Major Pi1 burst starts at 0703 UT (to right of 5040).



To investigate the degree to which a simple electrojet model as described above in fact represents station data, and since the necking region was near the station at Island Lake (ISLL), a reconstructed magnetogram was made for that station as shown in Figure 9.11. It is immediately seen that the X fit is excellent while the more sensitive Z component is less well represented by the simple electrojet model. A jump to nearly 200 nT in the Z component shortly after onset time (i.e. at 7.08 UT) has caused the fitting routine to respond and fit the Z much better at that time than at other times close to onset. Noting this allows a partial explanation of the narrowing of the electrojet inferred from the simple electrojet model. Clearly the strong currents flowing at onset have been concentrated somewhat south of Island Lake. An increase in the Z component at the next station north, Gillam, is seen at 7.12 UT (data not shown), about 2.5 minute later. The model does not respond to this peak in the same manner as by then the overall current has strengthened and the increase in Z no longer dominates the fit. This peak may represent a secondary intensification near (south of) Gillam, an interpretation which is supported by examining the MPA emission data.

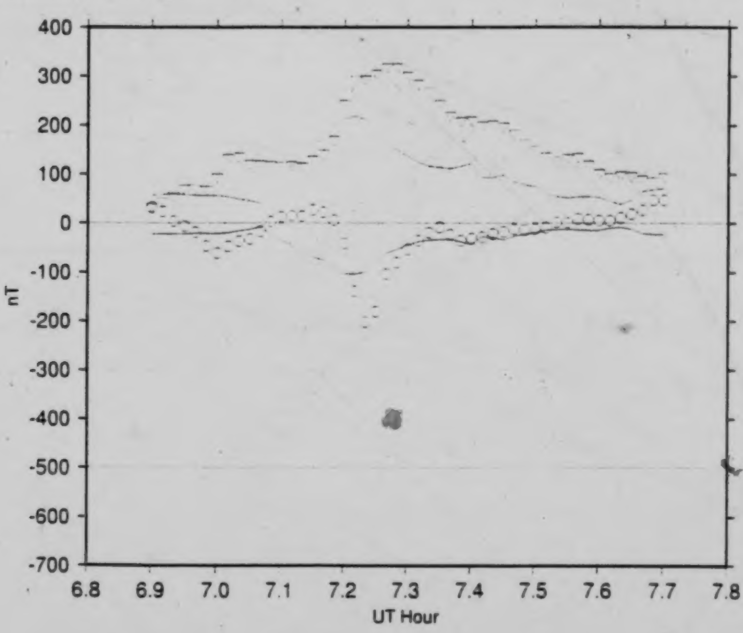
Figure 9.11 Data and reconstructed magnetogram for Island Lake on 8 August 1991. Circles represent observed X component (bottom), triangles observed Z component (top). Nearby lines are the corresponding modelling results.



Qualitative comparison of observed and modelled perturbations may also be done. A set of observed-modelled pairs may be constructed for each component. When this is done for the X component at Island Lake, it is found that for the period shown there is a Pearson r correlation coefficient of 0.9865 between the observed and modelled data. An actual linear fit between the observed and modelled pairs produces a slope of 1.0875, consistent with the model producing a result which is too negative for the (already negative) values of X. For the Z component, the value of r is 0.8265, not as good as for the X component. The 'sensitivity' of the Z component has been discussed several times and this is a further example of it. The slope for the Z correlation is only 0.6249, indicating that in general the model underpredicted the Z value.

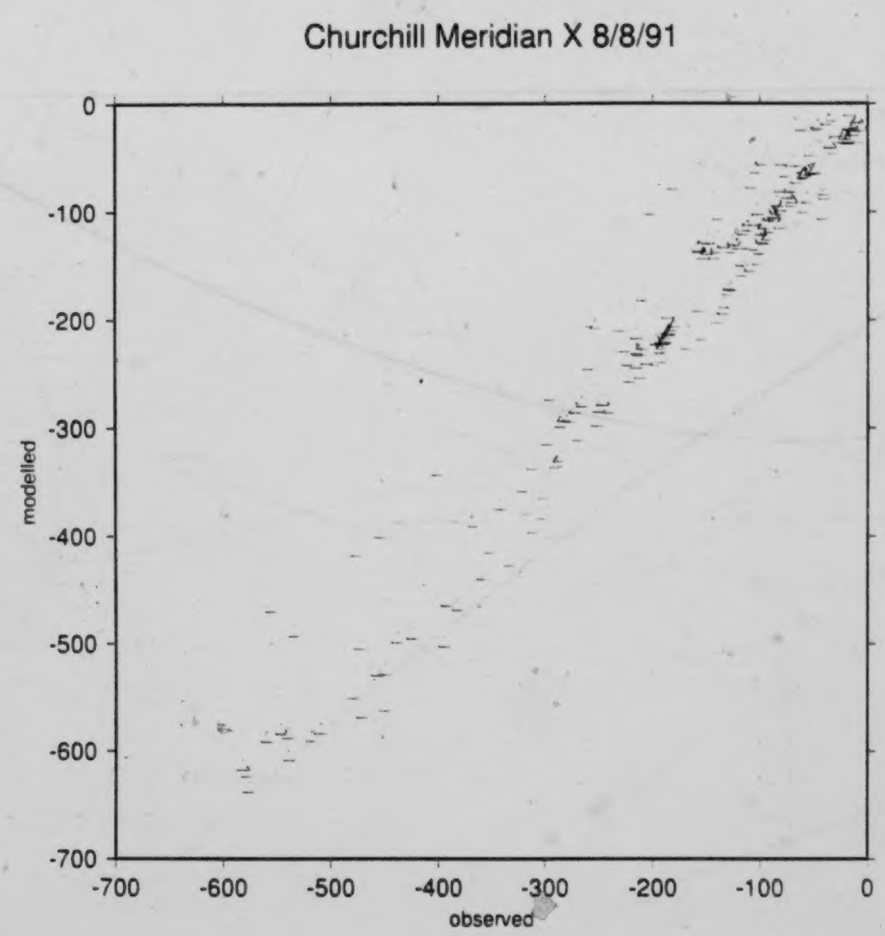
Inspection of more northerly stations in this rather southerly event showed agreement between data and model which was not as good as that at Island Lake. Figure 9.12 illustrates this fact for Fort Churchill (FCC). While the next station south, Gillam (data not shown), shows agreement almost as good as that at Island Lake, FCC shows only the most general agreement between modelling and observations. This is partly due to the fact that the perturbations there are smaller than at the two aforementioned stations, which thus dominate the modelling. It would certainly be possible to change station weighting and to improve the fit at FCC while not appreciably worsening it at the stations with larger perturbations. However, in a routine for general use, this type of 'tuning' is not felt to be appropriate. Quantitatively, the Pearson r coefficient between the X data and fit is only 0.5432, with a slope of 0.2776 confirming that not much of the X variation has in fact been properly modelled at this station. The corresponding values for Z are 0.7289 and 0.4730, indicating that Z has been better modelled here than has X. This result is the opposite of that found near the centre of the electrojet in that there, X is usually modelled better than is Z. However, it is understandable near the electrojet border where X is varying rapidly and Z less so. It is interesting to examine the data for time 7.08 UT at FCC. Changes are not large but the previously negative X component does approach zero then, so that the claim of electrojet narrowing is not negated by what this station shows. Similarly, there is a decrease in magnitude of the X component at Gillam before its onset-associated increase (data not shown). At both Gillam and FCC, these changes are on a longer timescale than the inferred electrojet narrowing. While there is nothing to contradict an overall narrowing of the electrojet at these stations, the interpretation above that most of the apparent narrowing is due to a localized intensification of current remains valid.

Figure 9.12 Data and reconstructed magnetogram for Fort Churchill on 8 August 1991. Symbols as in Figure 9.11.



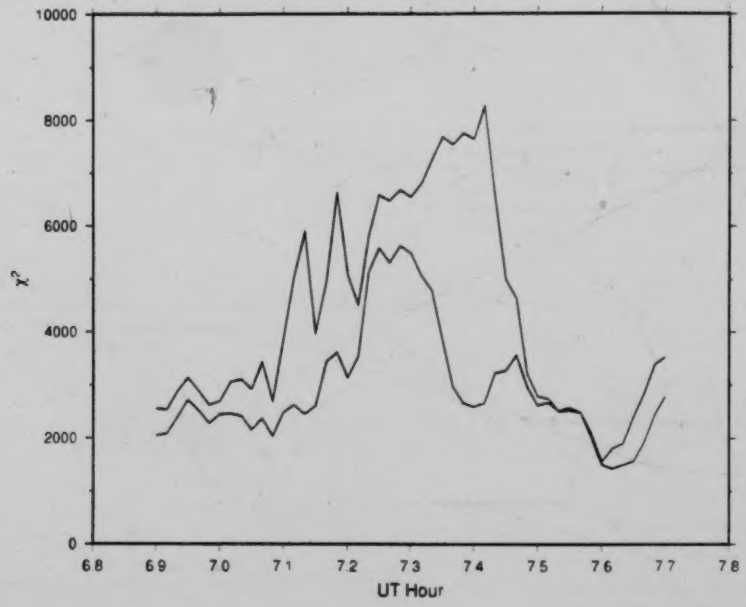
While lack of good modelling of an individual station may be felt to be a flaw in the attempt to use a simple electrojet model, it should be noted that the data from all seven stations used results in a correlation coefficient between X data and model of 0.9673, with a slope of 1.034. The scatter diagram of this relationship is shown in Figure 9.13. The figure represents all X data observed at all modelled stations through the event, paired with corresponding modelling results, and shows that there is a very strong relationship between these, a characteristic of a 'good' model. The square of the Pearson r correlation coefficient is a good indicator of the proportion of variation in one variable explained by another or by a model. In this case this 'coefficient of determination' is 0.9360, indicating that only about 6% of the variation in the data is not explained by the simple model. The overall representation of Z has a correlation coefficient of 0.9487, thus in principle representing 90% of the variation. The slope is only 0.7603: with most Z values being positive they are underrepresented by the model.

Figure 9.13 Relationship between observed and modelled X data on 8 August 1991.



As indicated above, it would be possible to tune station weightings to obtain a better data representation in this particular case, but that is not a useful thing to do here. What is more useful, after realizing that the simplest electrojet model, with only three parameters, accounts for upwards of 90% of the data variation, is to increase the number of parameters. This is suggested also by the bifurcation of luminosity distributions seen in MPA data. The number of parameters was doubled to six giving two electrojets. Given that there seems to be activity in the southern electrojet, the initial condition had overlapping electrojets with a common southern border. From this starting condition the routine generally separated the electrojets into a distinct northern and southern portion as will be discussed below. This results in a marginal increase in the correlation coefficient between data and model in X: it rises to 0.9706, but the slope becomes 0.9975. In Z, the improvement is more dramatic with the correlation coefficient rising to 0.9644 and a dramatic improvement in slope to 0.8548. The previous inability to represent the magnitude of the Z component has largely been overcome by the use of more parameters: more structure in the electrojet has resulted in a better fit to Z. The X component, varying little with latitude, is insensitive to this, but the more rapidly varying Z component is greatly improved with a model having smaller scale detail in the electrojet. Another aspect of fitting is the value of chi-squared attained. Figure 9.14 shows the improvement obtained, which is at certain times quite dramatic. It is less so immediately after 7.2 UT, the time at which FCC observes an steep decline in the X component. Inspection of the observed-modelled plot (not shown) for this station shows improvement at all times, but not dramatic improvement, and in particular the X decline after 7.2 UT is not strong enough in the modelling. This presumably reflects some relatively localized current distribution near the station, which cannot be adequately modelled if the station is given equal weight to all the others.

Figure 9.14 Value of chi-squared in two modelling runs of the Churchill meridian. Top curve is for a simple electrojet while bottom curve is for a two electrojet model.



There is no doubt that the two-component electrojet model has improved the fit as judged from its statistical properties and from the chi-squared attained. The results will now be examined critically to see what new information has been extracted. The two sets of electrojet boundaries are presented in Figure 9.15. The actual initial condition for this modelling run specified overlapping electrojets from 62.34° to 63.8° and from 62.34° to 66.8° (each with .547 MA of westward directed current). It is clearly seen that the solution is quite different from this initial condition and has enhanced the bifurcation of the electrojet. Comparison with Figure 9.6 reveals that the overall region containing electrojet flow is much the same as in the single electrojet case, but that the optimal solution with two electrojets is bifurcated, with a region of no current flow separating them at most times (despite this not being part of the initial condition). This bifurcation makes the pattern of electrojet flow correspond more closely to the luminosity in the meridian as detected by the MPA data (see Figure 9.4). Before 7 UT, the northward system shows southward motion as a coherent feature, much as shown by the more poleward arc in the MPA scans. The southern system is relatively fixed in latitude and there is a stationary emission region in a corresponding location (best seen in the Pinawa MPA). After 7 UT the poleward border of the northernmost system has expanded very much poleward and is not seen on the plots. This does not correspond to the behaviour of the northerly system in the MPA data but may reflect the ability of the model to represent weak currents by 'stretching out' latitudinally the region through which they flow. The net result is that current flow is effectively only near the southern border, not inconsistent with the 'necking down' seen near onset as discussed above. After onset at 7.08 UT, both current systems are clearly defined and are separated by a gap. In both the MPA data and the two-component model, this gap disappears through equatorward motion of the poleward system. The magnetic model indicates that the northernmost system becomes very broad, however, whereas the equatorward moving emission is a relatively narrow feature. Nevertheless, the overall morphology of the two-electrojet model appears to correspond well to that of the regions of emission throughout the event. This shows that the use of extra parameters in the magnetic model has given more information and that that information is consistent with that from other sources.

Figure 9.16 shows the current inferred for each electrojet in the two-electrojet model. The currents are shown as negative values, here indicating westward current flow. It may readily be verified that the sum of the currents shown is very close to the total current in the single electrojet model as shown in Figure 9.7, with maximum magnitude of about 0.7 MA near 7.2 UT. The magnitude of the current in the equatorward system increased steadily at onset time, and to a higher value than that in the poleward system. The latter increased also, but in more impulsive fashion and not as much. This is consistent with the conclusion that the onset initiated in the equatorward region.

Figure 9.15 Electrojet boundaries in a two-electrojet model for August 8 1991. Filled circles represent the more northerly electrojet.

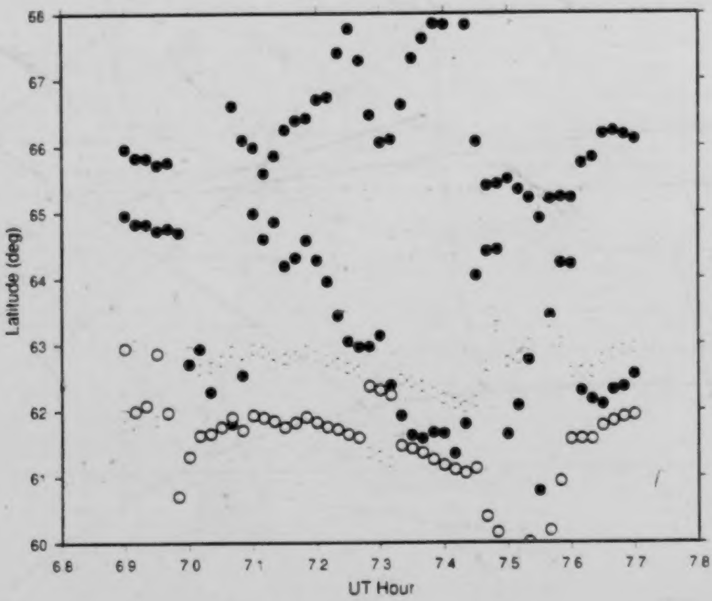
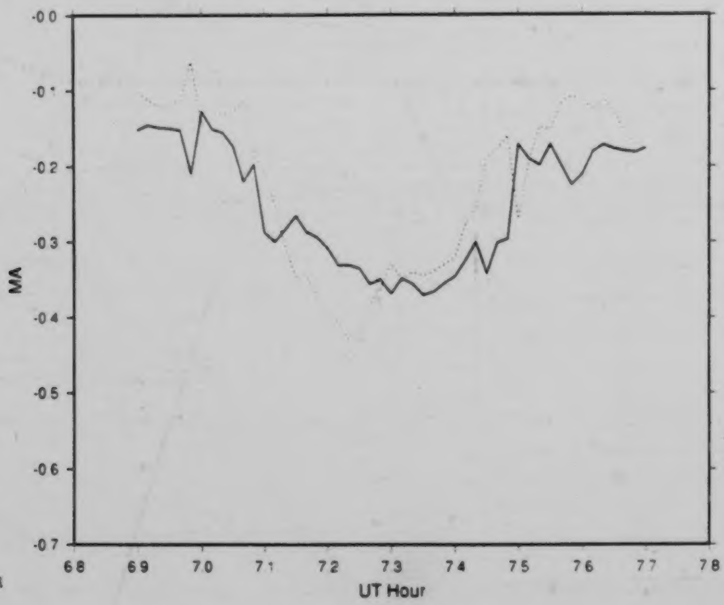


Figure 9.16 Current in each electrojet of the two-electrojet model for August 8 1991. Equatorward system marked by dotted line, poleward system by solid line.



This concludes discussion of this onset as modelled with Churchill meridian data. Simple electrojet models can give a very good reproduction of the behaviour of current flow over a meridian chain, in an automated fashion, and from poor starting conditions. Changing station weights would be expected to improve this ability further but would remove generality from the approach. In this specific case, good agreement was found between magnetic and optical data indicating colocation of electrojet and emission regions. The latter can be somewhat more filamentary than the former and this may be due to field-aligned current associated with the north-south Pedersen current. Such field-aligned currents would be associated only with the border of an electrojet. The meridian modelling indicates a narrowing of the electrojet if a single component electrojet is used. This is consistent with what was shown with a similar approach in Chapter 7. A two-electrojet model shows that there is bifurcation and that the more equatorward electrojet intensified at onset.

Meridian chain data suitable for modelling also became available for the Alaska stations which were in the high-latitude evening sector at the time of onset. The stations AMU, AVI, FYU, and TLK, whose parameters are given in Table 9.1, were used. Central to these was the standard station CMO, and north of them the Arctic coast station BRW. The latter had large Z component perturbations and could not be used in the electrojet modelling. These perturbations will be discussed in a wider context in a later section.

As seen in Figure 9.17, the evening sector current was not particularly intense, nor did it change much at onset time.

Figure 9.17 Current through the Alaska meridian on August 8 1991.

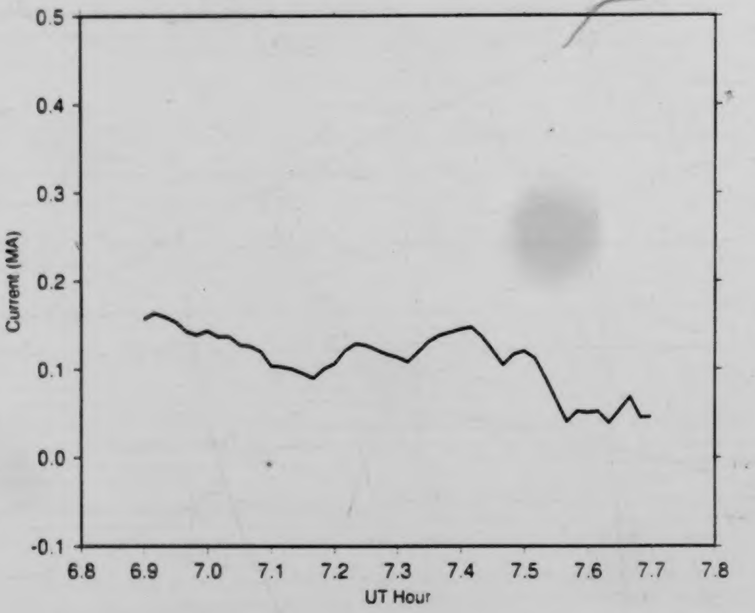
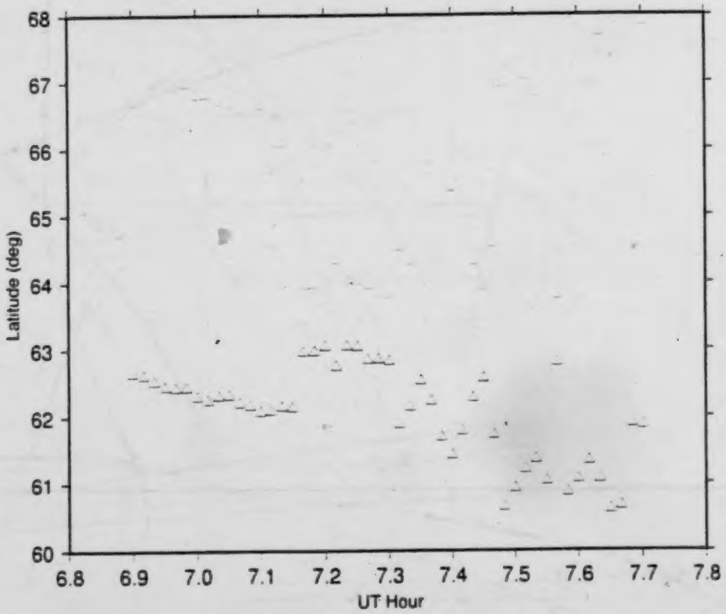


Figure 9.18 Electrojet borders in the Alaska sector on August 8 1991.



There is, however, a reaction to the onset in the form of a narrowing of the electrojet as shown in Figure 9.18. This narrowing is delayed with respect to the 7.08 UT onset, taking place at about 7.15 UT, that is, about 4 minutes later. Given that a rather strong substorm current wedge exists nearby at this time, its possible distant effects in the area of relatively weak eastward electrojet must be considered, and that will be done in the more complete modelling below. Here we will continue to interpret in terms of current systems which vary only in one dimension, that is, the latitude on the station chain meridian. In that case, it is useful to examine the latitude profiles directly to verify that the program has correctly indicated the electrojet width. Such profiles are shown in Figures 9.19 and 9.20, respectively for times 7.10 UT, immediately before the narrowing started, and 7.28 UT, while the narrow electrojet exists. It may be seen from both the X and Z components that the claim of narrowing is consistent in the context of single electrojet modelling, although the negative perturbations at the poleward border go unexplained in such a context.

A further interesting aspect of the electrojet borders as shown in Figure 9.18 is that after the period of narrowing, there is pulsational variation of the width of the electrojet during the time interval 7.3 to 7.5 UT, with period about 0.1 hour or 6 minutes (Pc 5 band). During this time interval, the overall current strength in the eastward electrojet decreases, so that after 7.5 UT it is essentially nonexistent. The large widths shown after 7.5 UT in the borders diagram are thus likely not very meaningful due to poor signal strength. Both of these claims may be verified by examining station data, and data from College and from Talkeetna 2.5° south of it are shown in Figures 9.21 and 9.22, respectively. Both stations show good fits, and during the period of the pulsations the X component peaks at the same time at each, indicating overhead current at both simultaneously and thus an overall widening of the electrojet. Although Talkeetna shows very clear pulsations in the Z component also, these are not obvious at College. At 62.3° centre dipole latitude, Talkeetna is nearer the centre of the electrojet at the times of interest and thus has more response in the Z component than does College which is closer to the northern border. There is also some evidence that more than one frequency may be present as there is a dip in the peak in X at Talkeetna which occurs at 7.4 UT. Another notable feature from these magnetograms is the decline in electrojet strength after 7.5 UT as already referred to. This is obvious in both magnetograms. The College data show negative Z starting at 7.2 UT, that is, at the time of electrojet narrowing. At 64.9°, College was prior to this time near the centre of the electrojet (see Figure 9.18). The narrowed electrojet is centred considerably south of the station, thus it has a decrease in X intensity accompanied by negative Z after the narrowing. It may be worth noting that the narrowed electrojet is entirely in the southern half of the latitude range occupied by the electrojet prior to its narrowing.

From the latitude profiles, it appears that electrojet modelling is appropriate (i.e. that local effects of the evening sector electrojet dominated the measured perturbations) in the Alaska sector at this time. Based on this, consistent results have been obtained from the modelling and from inspection of latitude profiles and individual station magnetograms. It could be concluded from these that narrowing of the evening sector eastward electrojet took place during this event and that the narrowed electrojet was in the equatorward part of the latitude range previously occupied by the electrojet. This change appears to have been the only one to immediately follow onset as the total current in the electrojet did not change appreciably after onset, until the decline of the substorm current wedge about 7.5 UT, at which time the eastward electrojet also decreased significantly in strength.

Figure 9.19 Latitude profile of Alaska chain stations at 7.10 UT August 10 1991. X component is indicated by filled circles for station data and nearby line for model results. Z component data is indicated by inverted triangles and nearby line is model result. Upright triangles indicate Y component data. Grey band along abscissa indicates latitudinal extent of electrojet.

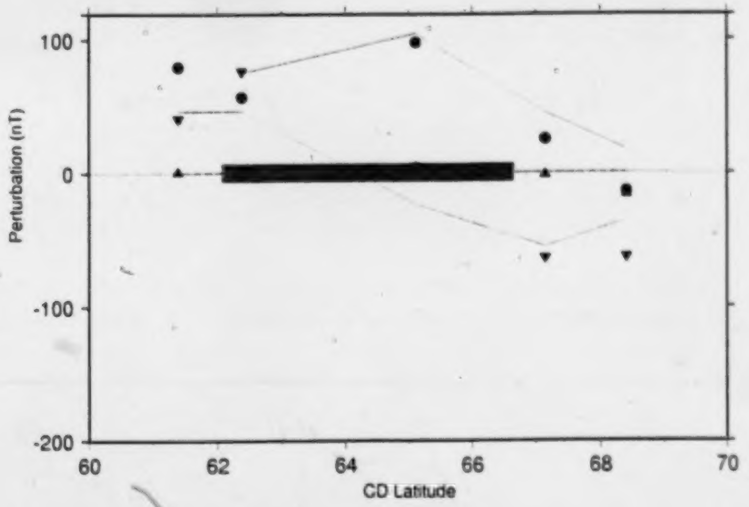


Figure 9.20 Latitude profile of Alaska chain stations at 7.10 UT August 10 1991. Symbols as in previous figure. More peaked X component data and positions of Z extrema indicate a narrower electrojet than in previous figure.

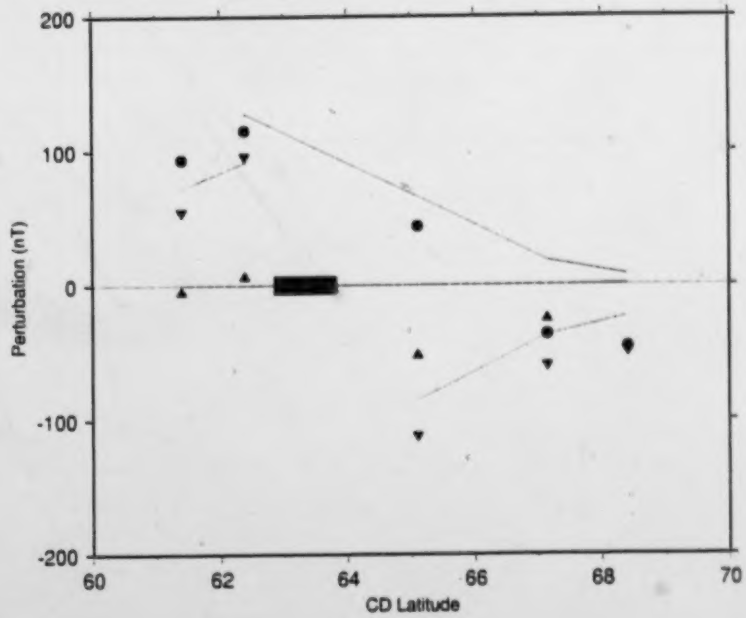


Figure 9.21 Observed and modelled perturbations at College on August 8 1991. Circles indicate X component data and nearby line represents model results. Inverted triangles indicate Z component data and nearby line indicates model results.

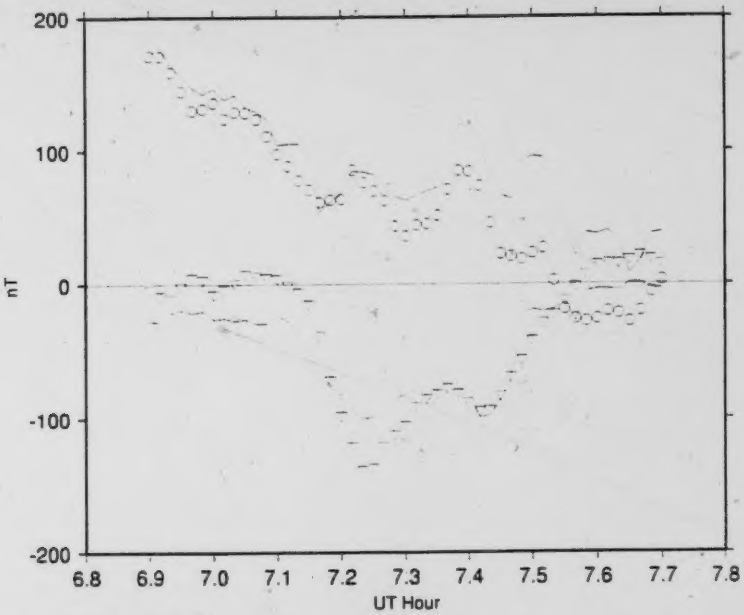
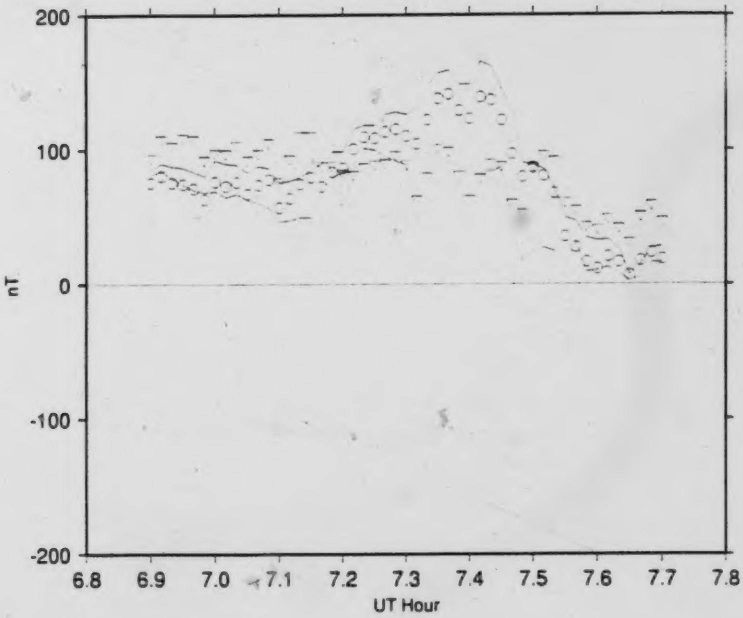
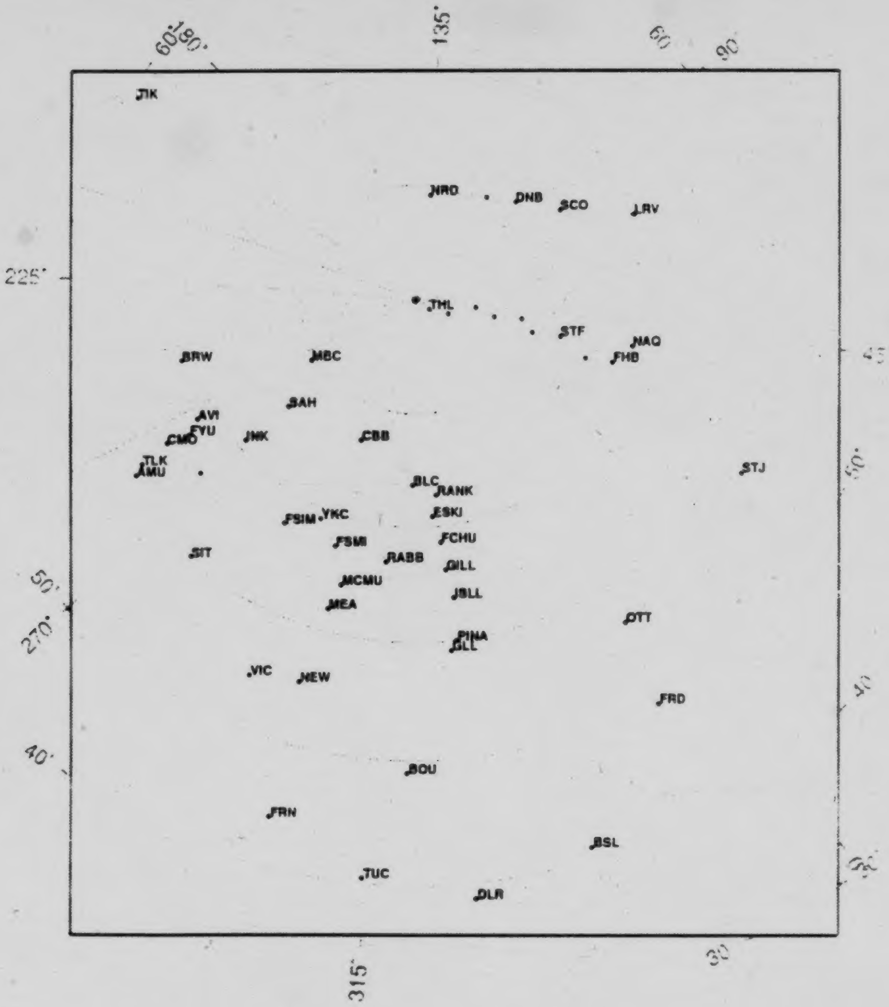


Figure 9.22 Observed and modelled perturbations at Talkeetna on August 8 1991. Symbols are as in previous figure.



It was intended to perform electrojet modelling on Greenland data which were obtained toward the end of this study. Examination of the data, however, showed that the electrojet was centred over the very most southern part of Greenland and that there would be uncertainty in determining anything about its southern border due to lack of stations, much as was the case in Chapter 8. Data from the station GHB was not available, meaning that only NAQ and FHB had much response to the electrojet. Stations at comparable magnetic latitude in eastern Greenland, along with Leirvogur (LRV) in Iceland, were available, but more suited to use in a regional modelling approach than with latitude profiles. For that reason, latitude profile studies of the morning sector westward electrojet were not performed independently and that electrojet was studied only with global modelling as described below.

Figure 9.23 Station key for modelling for August 8 1991. Named stations were used at some point but only North American stations and THL for 'wedge-only' case.


60°
180°
135°
60
90
225°
45
50°
270°
40°
40
30
315°
30
TIK
NRD
DNB
SCO
LRV
THL
STF
NAQ
FHB
BRW
MBC
SAH
AVI
CMO
FYU
INK
CBB
TLK
AMU
BLC
RANK
STJ
FSIM
YKC
ESKI
FSMI
FCHU
SIT
RABB
GILL
MCMU
ISLL
MEA
OTT
PINA
GLL
VIC
NEW
FRD
BOU
FRN
BSL
TUC
DLR

d. Wedge modelling results

Due to the symmetric placement of the current wedge over North America as deduced from the midlatitude perturbations, and the lack of obvious onset signatures at other global stations (data inspected but not shown), it was decided initially to attempt to explain the onset signatures in most of North America in terms only of the single three-dimensional current system of the substorm current wedge. The strengths and weaknesses of that approach will be discussed prior to more complete modelling in the next section. A station key is provided in Figure 9.23, including stations later used in global modelling.

Modelling for the pre-onset condition is presented as Figure 9.24. Modelling was done with station weightings high in the subauroral zone, three times lower in the auroral zone, and three times lower still in the polar cap and at College where it was known that the eastward electrojet was dominant. It is seen that at the polar cap stations the depicted wedge-like system did not produce perturbations capable of explaining those observed, a not unexpected result. Further, at College, the almost directly northward perturbation observed is a clear indication of strong eastward electrojet as detailed in the preceding section, which is not incorporated in the single wedge modelling. The best match at an auroral zone station is at Island Lake. At the next station south, the modelling at Pinawa places the electrojet over the station when it is not. Northward, the observed perturbation at Gillam is about twice as large as that given by the model, while further west the model's horizontal perturbation is larger than that observed at Rabbit Lake. Yet further west, there appears to be considerable westward horizontal perturbation which is not modelled adequately (at MCMU, FSMI, FSIM).

The conditions near onset time are modelled in Figure 9.25. Perturbations at Island Lake are again extremely well modelled while those at the next station north, Gillam, are not adequately represented. Once more, this is due to constraint from small perturbations at Rabbit Lake further west. The electrojet is seen to have narrowed considerably as previously indicated by the meridian chain modelling. Despite this, Pinawa had a modelled horizontal perturbation which is larger in magnitude than that observed. This discrepancy led to checking the data file, which was found to be in error for this one station. It was subsequently corrected for later use. Despite this error, the routine proved robust; the physical significance of the results was enough to point to an error in the input data. Since it was clear that a more sophisticated model would be run later, the simple wedge model was not rerun with corrected Pinawa data. Another mismatch, at Gillam, suggests that the westward currents should be located further north were they not constrained by Rabbit Lake, or that the northern border of the electrojet is not rectilinear as the model assumes. In this sense, little has been gained in the Churchill meridian by attempting a two-dimensional model. In general, with respect to the pre-onset condition, southward perturbations have about doubled across the North American auroral zone. At the western end of the activated region there remain significant westward perturbations which have not significantly changed in magnitude, and which are not accounted for by the model. The midlatitude perturbations are well modelled except near the ends of the wedge, where the influence of other current systems has clearly been felt. This applies at Sitka, where the wedge's equatorward horizontal perturbation is likely balanced by a poleward one from the eastward electrojet to form a slightly poleward perturbation. At St. John's similar considerations must also apply, although in this region, for the wedge modelling, it is likely also true that lack of stations prevents detailed modelling of the downward current associated with the current wedge.

The situation when the substorm current wedge was well developed is shown in Figure 9.26. Generally good agreement of all stations near the wedge except Pinawa may be seen. Gillam is now better represented as Rabbit Lake has also a negative X component. A current at the northern border of the electrojet appears to have progressed westward to cover the latter station. Gillam does, however, still not have the entire strength of its southward perturbation matched. The post-onset broadening of the electrojet is clear. Once more, there is a tendency for stations near the western end to show a westward component of the horizontal perturbations which is not reflected in the model. Also, College now had a considerably diminished positive X_m perturbation. Inspection of the magnetogram (see Figure 9.21) reveals that at this time the horizontal perturbations at that station were much decreased; after this time they remained small but a negative Z of over 100 nT magnitude developed.

Barrow, north of College, was not included in this 'single-wedge' modelling but does contain information about the evening sector eastward electrojet. At 0600 UT there was a nearly 300 nT positive X perturbation at Barrow, with very little Z, indicating that the electrojet was likely centred above the station. After 0600, the magnitude of the X component decreased steadily, becoming near-zero at about 0700. Simultaneously, the Z component (negative) increased in magnitude, reaching about -150 nT at 0700 and decreasing slightly in magnitude after the midnight sector onset. This behaviour, along with the peaking of the College X_m component much later than that at Barrow (at about 0650 UT), suggests that the evening sector electrojet moved southward prior to onset. This result should not be confused with the conclusions of the previous section about electrojet movement near onset time. Rather it may indicate equatorward expansion of the auroral oval preceding onset, rotation of the stations through a region in which the electrojet was dipping southward, or a combination of these effects.

To investigate the behaviour of the eastward electrojet and see the possible improvements at stations which were not well modelled by the single wedge, we now proceed to 'global' modelling, which is global in the sense that all major current systems are included, but has only hemispheric coverage by magnetic stations. The general result from the wedge modelling is that perturbations near the onset region can be well modelled, in general, by a single current wedge.

Figure 9.24 Single wedge modelling results at 0700.5 UT August 8, 1991, shortly before time of onset.

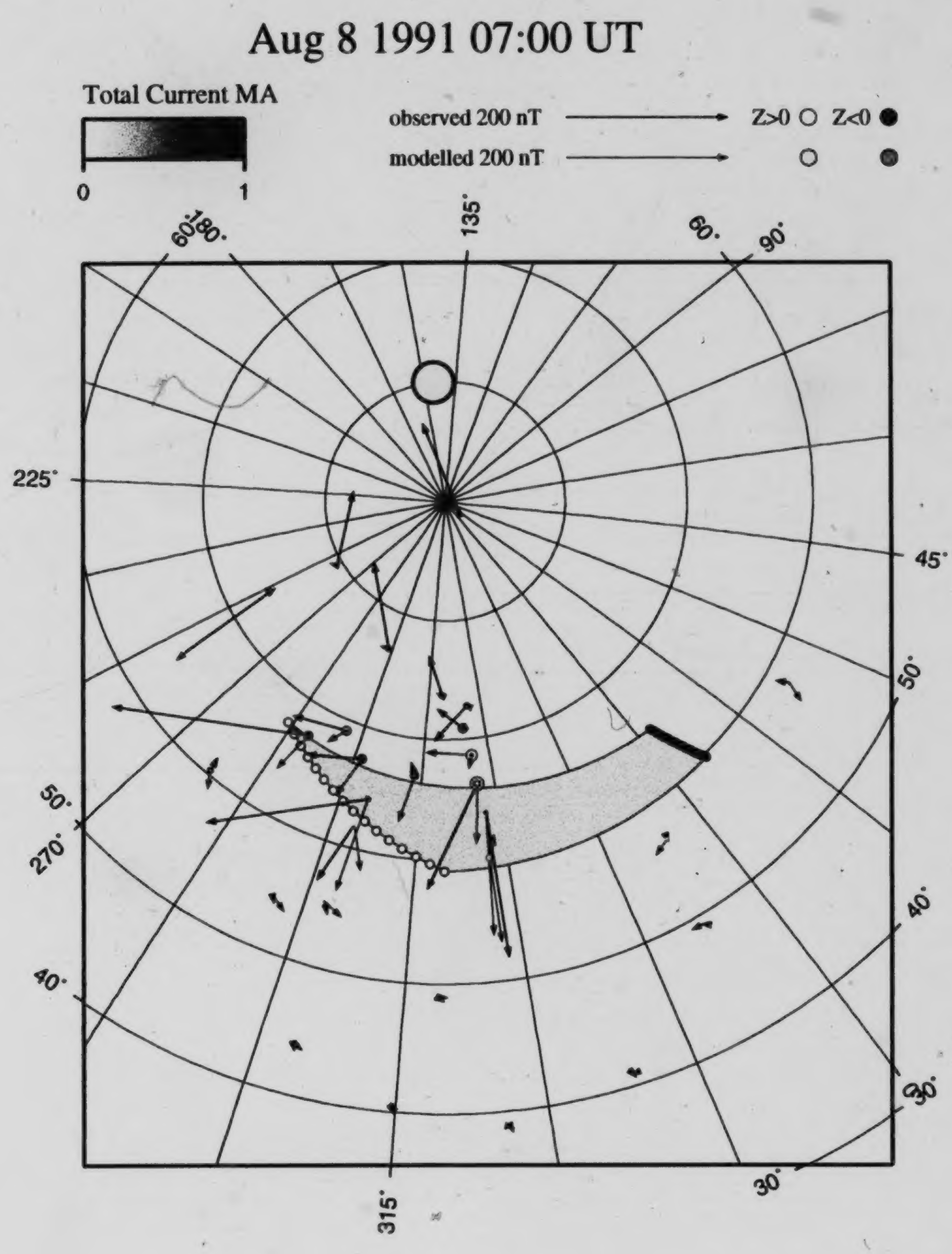
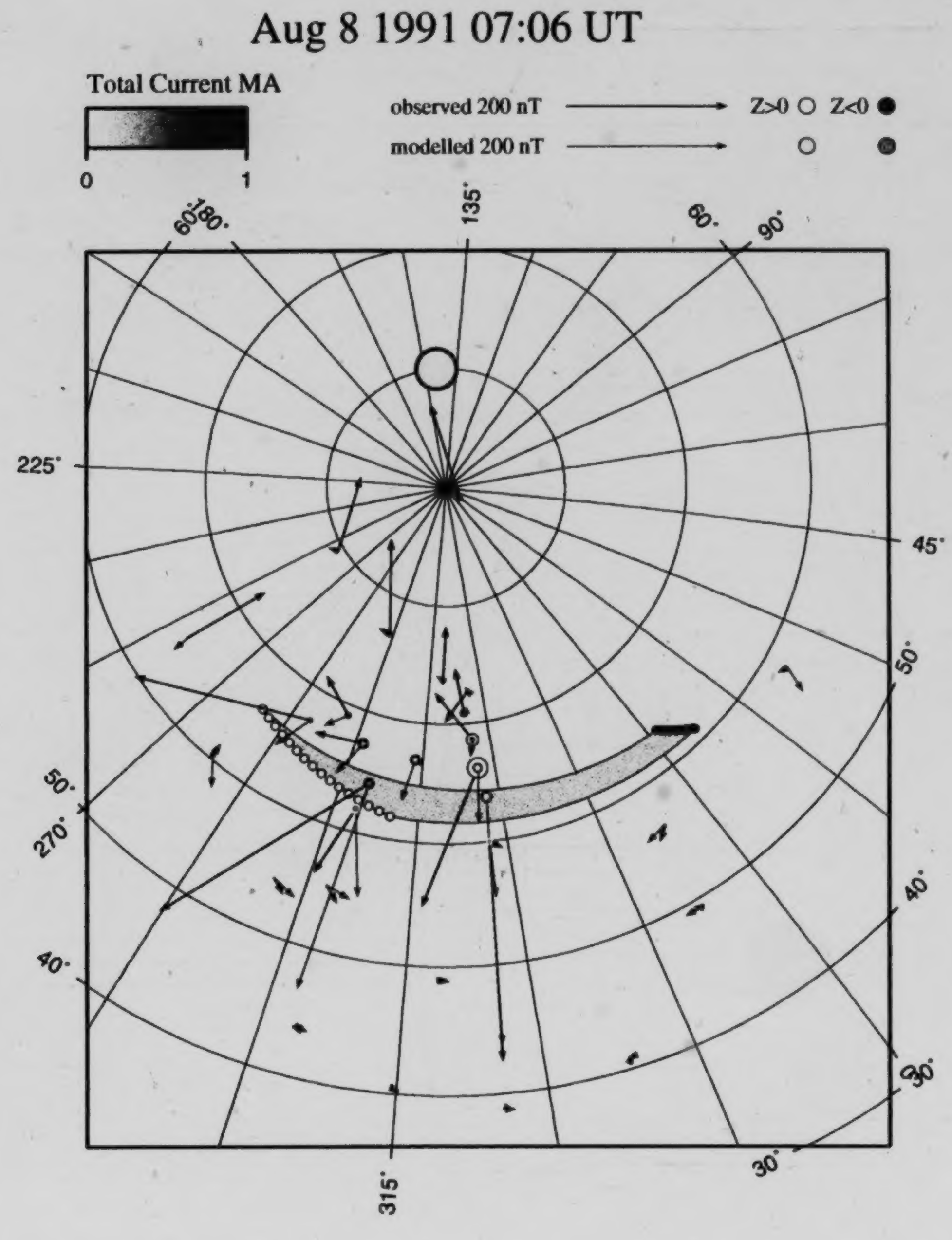
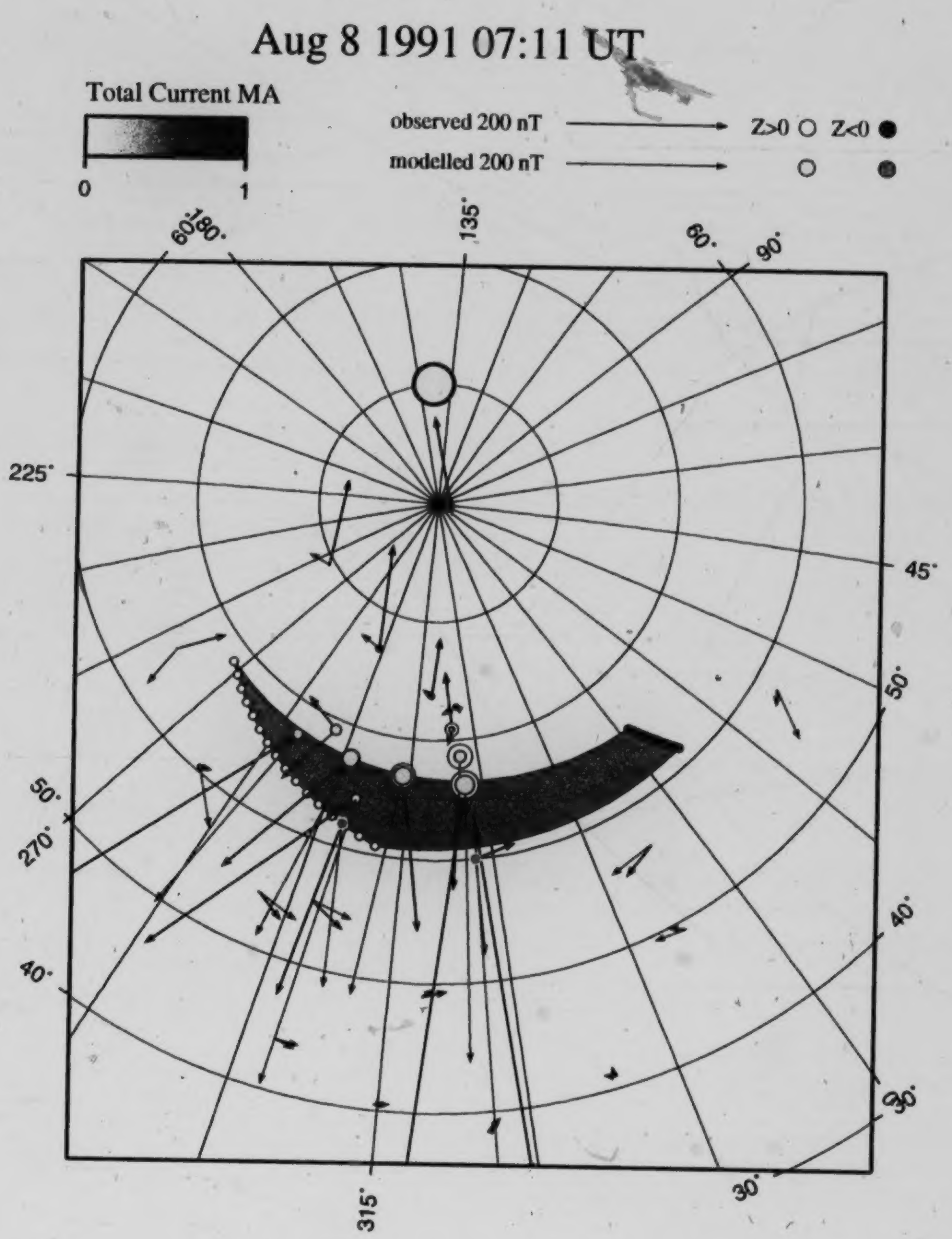

Aug 8 1991 07:00 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●
180°
60°
135°
60°
90°
225°
45°
50°
50°
270°
40°
40°
30°
30°
30°
315°

Figure 9.25 Single wedge modelling results at 0705.5 UT August 8, 1991. Observation arrow at ISLL (327.6,63.7) is obscured by model arrow.



Aug 8 1991 07:06 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○ Z<0 ●
180°
135°
90°
225°
45°
270°
315°
60°
60°
50°
50°
40°
40°
30°
30°

Figure 9.26 Single wedge modelling results at 0710.5 UT August 8, 1991.


Aug 8 1991 07:11 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●
60°
180°
135°
60°
90°
225°
45°
50°
270°
50°
40°
40°
30°
30°
315°

e. Global modelling results

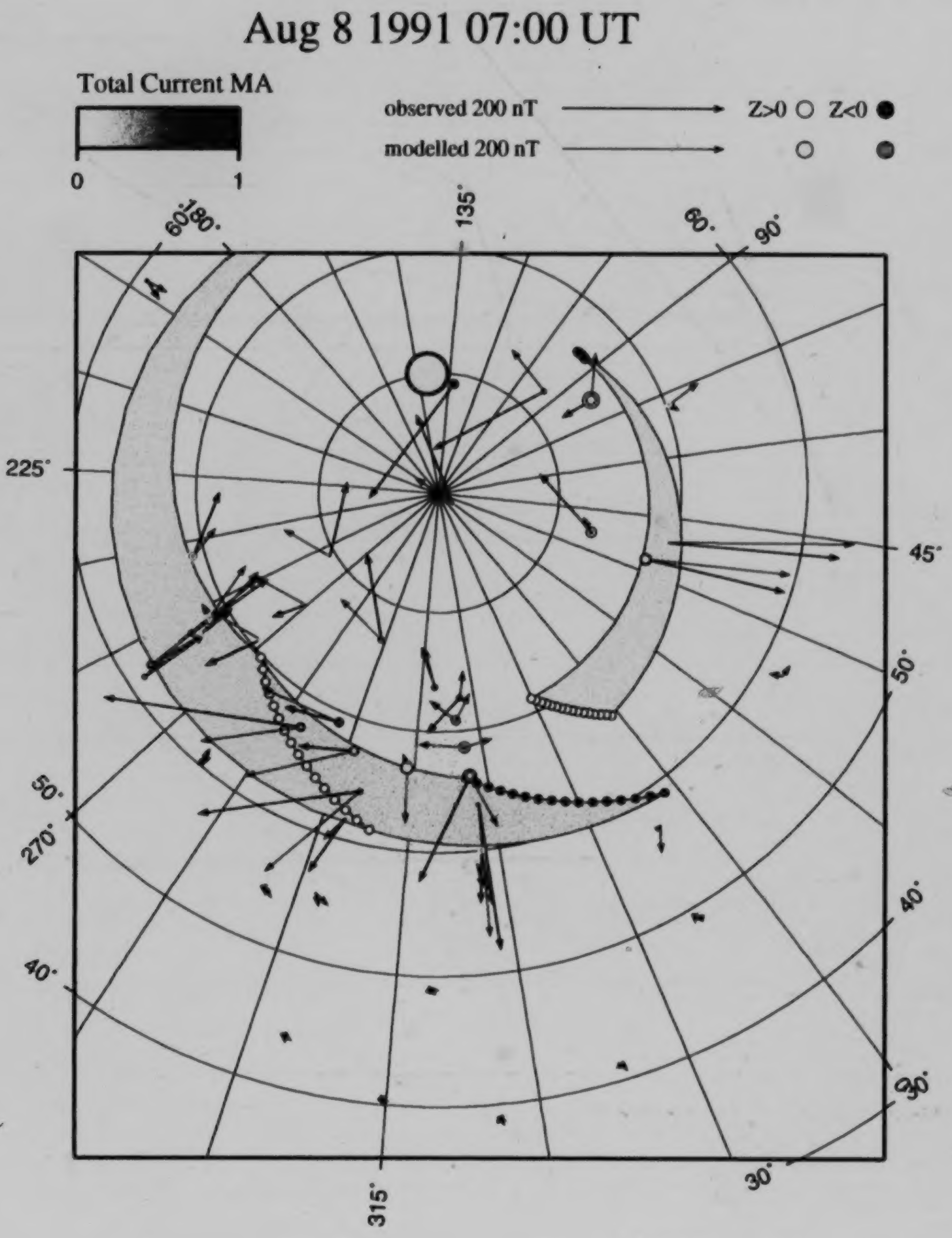
The single current wedge gave a relatively good representation of the currents over central North America, but there were indications that other current systems might also be important. Thus, an attempt was next made to study global or at least nightside electrojet development at onset time. In addition to the data sets cited above, recently obtained data from Greenland and Iceland were used in modelling, and European data were inspected. The one readily available Russian station, Tixie, does not contain much information of use in the modelling as it appears to be south of the auroral oval. This appears to also be the case for the EISCAT array in Scandinavia. While Tixie was used, the EISCAT data appeared redundant, as did northern Greenland data which did not have obvious substorm signatures. The final set of stations used is shown in Figure 9.23.

In addition to seeking a picture of current systems over a larger region, an attempt is made to explain perturbations not well accounted for by the simple electrojet model. These included primarily northward perturbations north of the midnight sector electrojet, and the westward deviations of the horizontal components as observed near the western end of the current wedge. Initial modelling runs included three current systems, those of the wedge much as described and presented above, the evening sector eastward electrojet, and the morning sector westward electrojet. The primarily toroidal currents of these three systems were found to allow good reproduction of most auroral zone and subauroral perturbations, but not of those in the polar cap. While these perturbations did not appear to be as organized as, for example, those during the CDAW 9B event of Chapter 6, a similar approach was taken in attempting to represent them as the distant effects of Region 1 and 2 field-aligned currents. The overall model system was modified to include north-south current (loosely this would be the Region 1/2 currents with associated ionospheric closure) in the same basic regions as the better-constrained toroidal current flows. The results of these full models will be compared to those from single wedge models below. The overall result was that polar cap perturbations were not notably well modelled through inclusion of north-south currents. The Chapter 6 conclusion that these currents dramatically improved ability to model the polar cap was obtained under equinoctial conditions, whereas at the summer date of this study, the polar cap was almost totally sunlit. It therefore cannot be excluded that the polar cap perturbations were dominated by local currents, rather than by the distant effect of field-aligned currents in the auroral oval as was the case in Chapter 6. A particular example of this will be seen in that the Hall-dominated DPY current system [Friis-Christensen, 1983] was obviously present as indicated by polar cap perturbations.

The colocation of the toroidal and poloidal currents was enforced through an algorithm which placed the longitudinal boundary of north-south current flow in the middle of sheets of field-aligned current delimiting the toroidal flow. Thus each north-south current had only its current as a completely free variable. A *de facto* Harang discontinuity was also imposed by requiring the field aligned current sheets at the western end of the wedge to have the same location as that of the sheet at the eastern end of the evening sector eastward electrojet. In a region of far fewer stations, the ends of the morning sector westward electrojet and of the nearby current wedge were left unconstrained. In the solutions they thus generally indicated upward and downward current near one another but not coincident. It should be recalled with reference to earlier modelling results that this may give an appearance of upward and downward current in the same general region, but that this is not actually implied by that configuration. Generally, the current wedge will be stronger than the morning sector electrojet, and the net current will be a downward one corresponding to the difference in current strength. For example, if a 0.1 MA westward electrojet is shown flowing upward near the downward current of a 0.5 MA current wedge, then what is physically happening is that there is 0.4 MA of downward flow into the wedge, and flow of the entire electrojet into that region also. On the diagrams used here, there would indeed be both upward and downward current nearly collocated, but this is an artifact of the model and indeed language used. In cases where the two currents are shown separated by a considerable distance, it is most likely that they are indeed joined, but this cannot be shown one way or the other with the data coverage available.

Figure 9.27 shows the pre-onset conditions with full global modelling and may be compared to the single wedge results of Figure 9.24. It can be noted that the electrojet in the midnight sector has its latitudinal borders in the about same place as in the single wedge result, that is from approximately Pinawa (59.9°) in the south to the Gillam-Rabbit Lake latitude (about 66°), for a total width of about 6 degrees. Meridian modelling at this time (see Figure 6.6) agrees on the northern border but places the southern border about 2 degrees further north. In fact, the global modelling has the southern border slightly further north than does the single wedge attempt, and this is likely due to the inclusion of Glenlea, slightly south of Pinawa, adding more weight to the southern region. Close inspection shows that the stations south of Gillam in the Churchill meridian are in fact extremely well represented by the global model. The addition of the evening sector eastward electrojet has allowed Alaskan stations to be well modelled in general, with Barrow (a low weight station) likely suggesting that the electrojet should be slightly further north. Tixie is relatively well represented and was given a heavier weight due to representing perturbations in a large region. The region near the interface of eastward and westward currents (Harang discontinuity) is well represented in the modelling. The considerable westward tilt of the horizontal perturbations has been better modelled than in the corresponding 'wedge-only' model, likely mainly through the strengthening of total upward current in the Harang region, which now has both electrojets feeding a field-aligned current there. The morning sector electrojet crosses southern Greenland and the model represents it well there. As has already been discussed, the gap between it and the midnight sector electrojet is not likely real, and the similarity of total current in these systems indicates that they are one system. However, the station gap in northeastern Canada does not allow this to be verified.

Figure 9.27 Global modelling results at 0700.5 UT August 8, 1991.


Aug 8 1991 07:00 UT
Total Current MA
0
1
observed 200 nT
Z>0 ○ Z<0 ●
modelled 200 nT
○ ●
135°
180°
60°
60°
90°
225°
45°
50°
50°
270°
40°
40°
30°
30°
315°

A final note about the pre-onset situation shown in Figure 9.27 is that the DPY current system is clearly active as seen in the noon sector stations, which with their northward perturbations indicate a high-latitude eastward electrojet. The model does not represent these perturbations well and they are likely associated with a local Hall current [Friis-Christensen, 1983]. DPY effects are not uncommon under summer conditions and they are associated with the Y component of the IMF [Rostoker, 1980]. IMF data is not available to confirm this connection for this event, and it was felt that adding this system would not add to the interpretive value of the present modelling effort.

Figure 9.28 shows the situation close to onset time and may be compared to Figure 9.25 from the wedge only model. The narrowing of the electrojets is immediately obvious. This may not be entirely appropriate in the Alaskan sector where AMU, the most southerly station, has larger perturbation northward than does the model. In fact, the meridian chain modelling indicated that narrowing in that sector had not taken place at 7.09 UT but only set in at about 7.15 UT. The narrowing in the midnight sector and the coupled boundaries of eastward and westward electrojets appear to have forced a premature narrowing in Alaska. It is also immediately clear in comparing to the previous figure that the narrowing, at least in the midnight sector, has primarily been through a contraction of the electrojet into the southern part of the region which it formerly occupied. In the Harang region, stations are once more fairly well represented. From stations INK and FSIM, it may be inferred that there is a northerly spur of westward current. This would account for the negative X_M perturbation at the former and could strengthen the horizontal perturbation at the latter. There and at MCMU, such a northerly trend of the electrojet would rotate the horizontal perturbation to improve the fit (i.e. clockwise). This spur may be inferred to be small, however, compared to that found in Chapter 6, which necessitated an entire new current system due to its size and inferred strength. This diversion of current to the north, not included in the model, may also explain the poor match at Sitka (SIT), which shows signs of nearby upward field-aligned current through its westward modelled perturbation. This is in fact not observed. Other subauroral and midlatitude stations in general have an excellent fit west of 330°. Ottawa does not show a good fit, having a southwestward perturbation which is too strong in the model. This is likely due to the gap between the morning and midnight sector electrojets, the upward FAC from the former being too far north to offset sufficiently the perturbation from the downward FAC of the latter. Perturbations in the southern part of the Churchill meridian, under the wedge, are extremely well matched. Gillam, near the northern border, does not show enough southward perturbation, likely indicating an enhanced northern border current. Churchill itself shows a positive perturbation which is not well modelled, but the more northerly stations in the chain have an improved agreement now that the model includes Region 1/2 currents. The morning sector electrojet appears relatively unchanged from five minutes before, as do DPY system and polar cap perturbations.

Figure 9.28 Global modelling results at 0705.5 UT August 8, 1991.

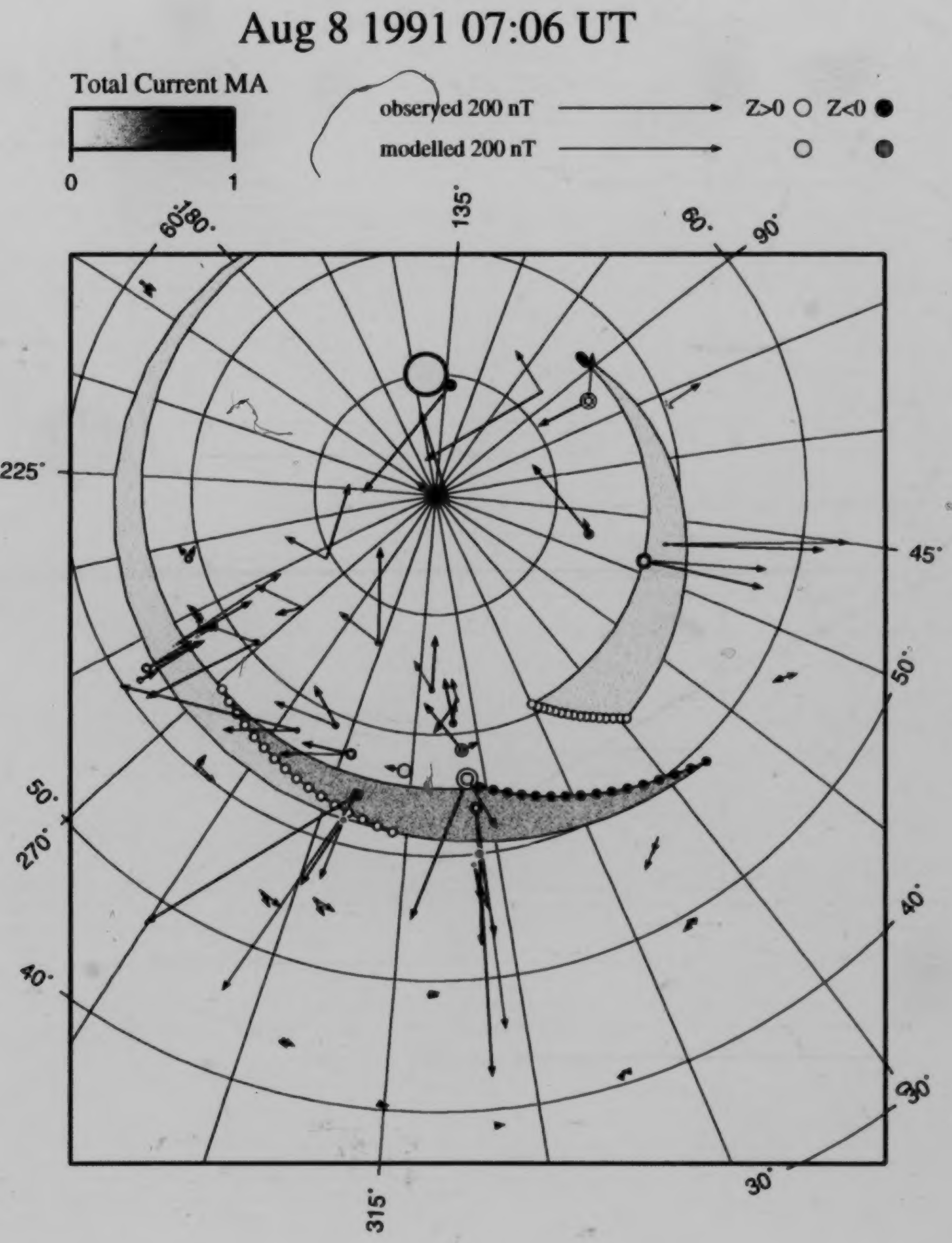
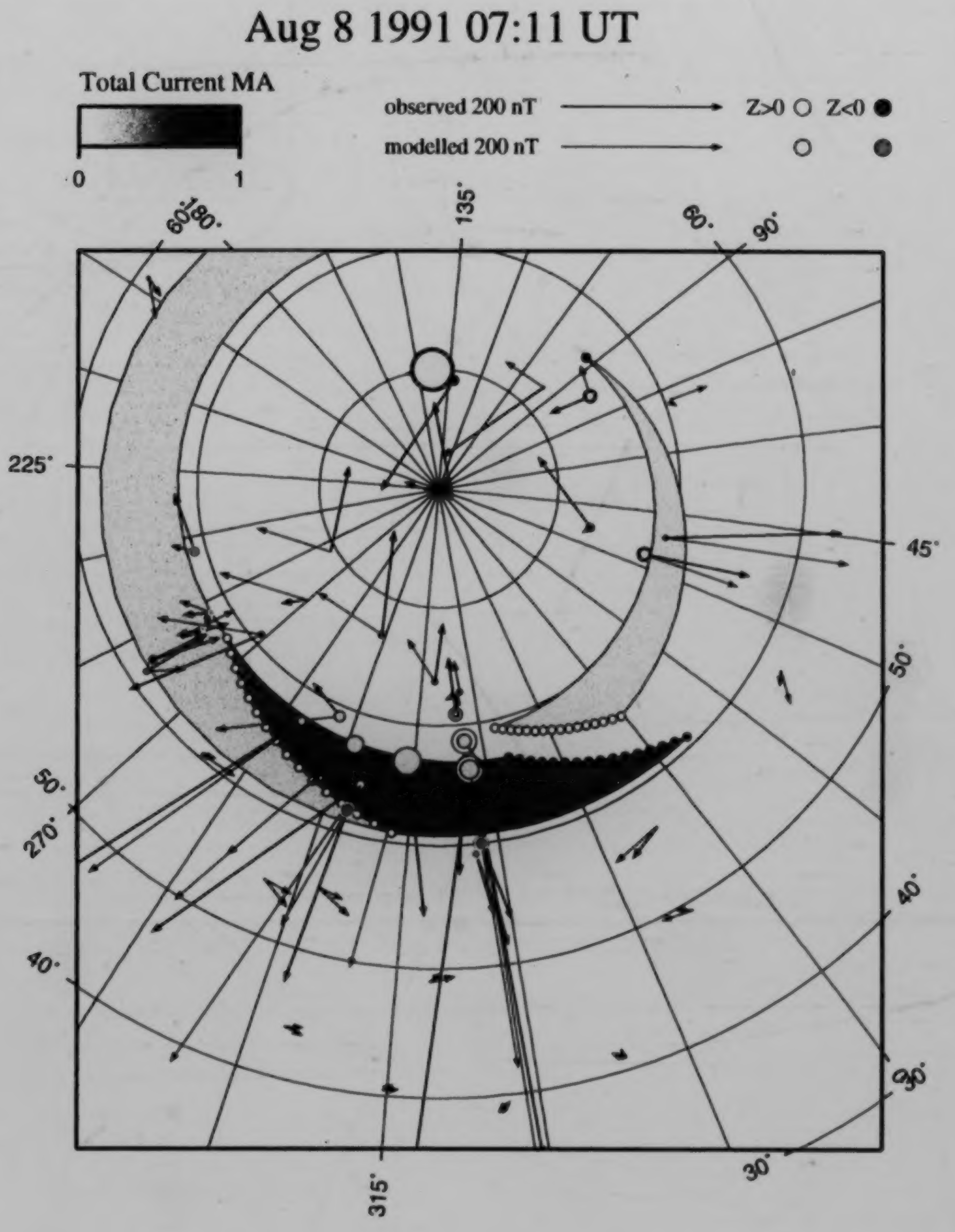

Aug 8 1991 07:06 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●
135°
60°
180°
60°
90°
225°
45°
50°
50°
270°
40°
40°
315°
30°
30°

Figure 9.29 Global modelling results at 0710.5 UT August 8, 1991.



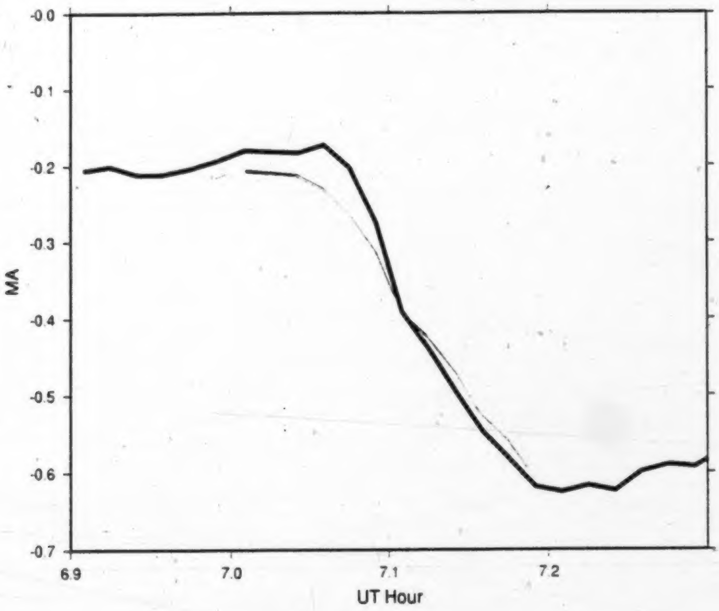
Aug 8 1991 07:11 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●
60°
180°
135°
60°
90°
225°
45°
50°
50°
270°
40°
40°
30°
30°
315°

Figure 9.29 shows the situation once the substorm current wedge had come to full strength. As concluded previously, the midnight sector electrojet has widened again by this time and this is clearly shown in the figure. However, examination of Figure 9.18 shows that this is a time when the evening sector electrojet was inferred to have narrowed. Due primarily to the linking of currents at the Harang discontinuity, this has not taken place in the model. It is clear in taking the global view that the westward electrojet intruded north of the evening sector electrojet as indicated by the strong southward component of observed field at Inuvik (INK) and some southward field at AVI. The model has to some extent indicated this intrusion through a northwestern motion of the northwest border of the wedge current, and the perturbations at the stations discussed are in fact relatively well represented in the model. Figure 9.20, presented to justify the conclusion that the evening sector electrojet had narrowed, shows negative perturbations at its northernmost stations (FYU and AVI; the former somewhat hard to see on Figure 9.29). These are now shown to be due to the intrusion of the westward electrojet related to the wedge, as was suspected but not provable earlier.

Remaining aspects of the modelling include a good agreement of the model perturbations with observation at most subauroral stations. In eastern North America, the perturbation at Ottawa is in good agreement while the observed increase at Fredericksburg appears less than that called for by the model. The morning sector electrojet through Greenland appears little changed although its terminus has come closer to the start region of the wedge, likely explaining the improvement at Ottawa. Perturbations associated with DPY also appear to be little affected by the midnight sector onset, a statement which applies equally to other polar cap perturbations. Despite the inclusion of north-south currents in the model, these continue not to be well modelled, although comparison of Figures 9.24 through 9.26 with the corresponding Figures 9.27 through 9.29 will show that addition of these currents did significantly increase the magnitude of horizontal perturbations in the polar cap.

As a check on the consistency of the global and meridian profile models, it is useful to compare the currents crossing the Churchill meridian in the two cases. Figure 9.30 makes this comparison. It is seen that there is excellent agreement between the two results. The influence of the small perturbations at Rabbit Lake, alluded to above in regard to single-wedge modelling, have likely been negated by having enough other stations that its weight becomes too small to influence the global modelling results (much as it does not influence the meridian chain results since it is not included in the meridian). The small perturbations at Rabbit Lake, and the relatively poor representation of Churchill in both the meridian chain results and global modelling (inspect Figures 9.27 through 9.29) indicate that structure exists at the northern border of the electrojet.

Figure 9.30 Comparison of eastward current through Churchill meridian from meridian and global modelling. Heavy line is result of meridian model.



f. Summary

This study has reinforced results presented previously about narrowing of electrojets at or shortly after onset time. The event studied initially appeared to be relatively isolated but as more data were examined, more activity was found. This is perhaps not unexpected given the low latitude of the onset which prompted selection of this time period for study. Selection on that basis alone also resulted in study of a period without solar wind data, so that conclusions about triggering through solar wind effects are not possible. Conclusions about auroral electrodynamic effects include:

1) the pre-existing westward electrojet in the near-midnight sector narrowed at onset time in a fashion consistent with concentration (possibly as narrow filaments) of current flow in its southern half. This is verified by meridian chain modelling, by a single wedge model, and by a global current model

2) the pre-existing eastward electrojet in the evening sector appeared to narrow shortly after onset time, its latitudinal centroid moving southward. Although the southward motion is real, much of the apparent narrowing is an artifact of intrusion of the westward electrojet at the poleward border

3) the midnight sector electrojet intensified to about 0.7 MA in becoming the current wedge electrojet

4) the evening sector eastward electrojet did not appreciably increase its total current at or near onset time

5) Pc 5 band pulsational activity in the evening sector electrojet could be well interpreted in terms of width changes of that electrojet

6) good correlation of the electrojet borders in the wedge region with optical emissions including bifurcations

7) poleward expansion in the onset region to a limited extent.

In addition, some quantitative aspects of the modelling were investigated. The conclusions about modelling included:

8) good representation of the Churchill meridian electrojet using only three parameters (total current and electrojet borders)

9) better representation of the meridian chain results using six parameters (two sets of electrojets)

10) good correlation between electrojet borders as determined by the modelling and as inferred by inspection of luminosity profiles, and coincidence of these borders with luminosity maxima

11) likely existence of structure at the northern electrojet border inferred through global modelling.

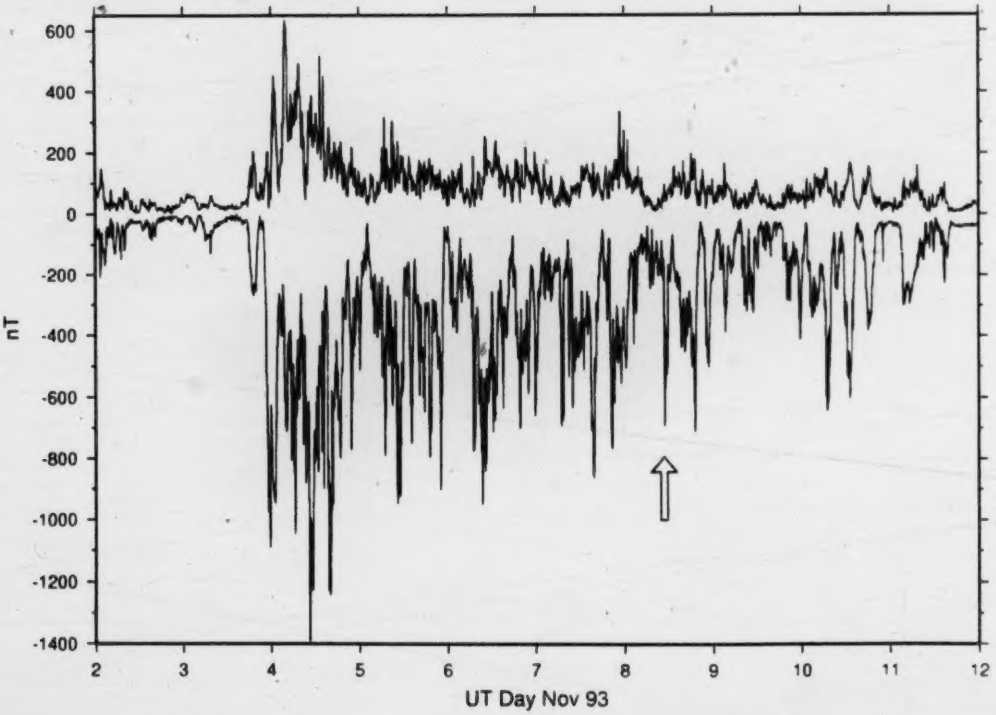
Chapter 10. GEM Event - November 8, 1993

A storm period beginning on November 3 1993 and lasting at least until November 10 has been selected as a target of detailed investigation in the Geospace Environment Modelling program (GEM). It is expected that a series of papers on the storm will be published in early 1998 in the Journal of Geophysical Research, with the initial onsets (during the first two days) having received the most attention by researchers to date. Here a substorm near the end of the storm period has been selected based on its clear AE signature and on the availability of IMP-8 interplanetary data. As part of a larger study, this event is characterized by good data availability with over one hundred stations being usable for northern hemisphere modelling routines and with many southern hemisphere stations also available for conjugate use. Available global five minute-averaged data will be used in this chapter, where the emphasis will be on global modelling using as many stations as possible.

a. Overview

The GEM storm is obvious on a plot of auroral electrojet indices for early November 1993 (Figure 10.1). Frequent substorms cause an impulsive nature in the AL index in particular. One of the few isolated substorms took place on 8 November.

Figure 10.1 AL and AU indices (bottom and top curves, respectively) on November 2 to 11, 1993. Many negative AL excursions correspond to substorms; that at day 8.5 (marked by arrow) is studied here. Computed by Y. Kamide from data at 89 stations.



The selected substorm has several hours of lower activity preceding it. It is also only in this latter part of the storm period that IMF data became available. Since the event occurs in a storm period which is the object of a major study, the events of that period will be briefly summarized here. Much of the following information is abstracted from a GEM project internet site[1] [D. Knipp and contributing scientists, 1996].

The late phase of Solar Cycle 26 was characterized by recurrent geomagnetic storms and the early November 1993 storm was the fourth recurrence out of eight associated with a long-lived coronal hole and high speed stream [Crooker and McAllister, 1997]. Such recurrences, as discussed in Section 1.1, are spaced in time by 27 days, the synodic rotation period of the Sun. In each of the five recurrences, enhancements of 6.0-7.8 MeV magnetospheric electrons were observed by LANL geosynchronous satellites. In this case, Yohkoh X-ray imagery of the Sun clearly showed a coronal hole from 30 October onward. Neither of these aspects is discussed further here.

In the pre-storm period, late November 2 to the early UT hours of November 3, magnetic activity was low (K_p index in 0 to 1- range). On November 3 at about 0600 UT there was a sudden increase in solar wind plasma density, sufficient to allow IMP-8 to sense solar wind plasma characteristics while apparently remaining in the magnetosheath. Solar wind velocity remained moderate at approximately 275 km/s and cross polar cap voltages were modelled (using AMIE) to be a modest 30 kV. At roughly 0700 UT the D_{st} initial phase began with an increase in D_{st} to about 30 nT. At 1720 UT November 3, a Sudden Commencement (SC) was observed and D_{st} began to decline. At about the same time as AL dropped precipitously, D_{st} rapidly declined to less than -100 nT. These classic signatures of a magnetic storm are shown in Figure 2.3, and the depressed values of D_{st}, characteristic of the storm main phase, lasted over one week. Late on November 3, the solar wind disturbance responsible for the storm contained an IMF which fluctuated with some southward periods until it turned strongly (-15 nT) southward between 2200 and 2300 UT and even attained 35 nT southward in the following hour. Extreme density enhancements (to nearly 100 per cm^3 or 10^8 m^{-3} number density) were also observed during this period. Major nightside X component bays were observed at many ground stations.

In the early UT hours of November 4, the magnetopause was observed to be inside geosynchronous orbit, further evidence of the extreme solar wind pressure (recall Figure 1.4 and accompanying discussion), and the IMF continued to be strongly southward. The activity index Kp maximized at 6- and Dst minimized at -105 nT. At about 0515 UT GEOTAIL observed a solar wind speed increase to a maximum of 700 km/s and the speed remained at this high level for the rest of the day. Preliminary AMIE results indicate cross-polar cap voltages in excess of 160 kV.

---

[1] http://igpp.ucla.edu/gem/event_nov93/newbackground.html; entitled "Background for November 1993 Space Weather Event Study"

Generally high levels of activity continued throughout the next several days, as indicated by the 89-station auroral electrojet (AE) indices [B. Emery, personal communication 1996] of Figure 10.1, but only on November 8 did IMP-8 emerge from the magnetosphere into the dawnside solar wind. It recorded hourly average velocities in excess of 600 km/s which declined steadily to more typical values over the following four days. Particle densities were back to typical values of 5 to 10 per cm^3, and the magnetic fields showed typical variations on time scales of tens of minutes, including several north-south transitions near the time of the onset studied here. The period studied in detail here is from 1000 UT to 1200 UT on November 8. The substorm in question may thus be studied with solar wind data available. Not until November 11 did D_{st} recover to near-prestorm values, and other activity indicators also decline in magnitude. The event under study thus took place under storm conditions during the declining phase of the storm.

b. Chronological Summary

Against the backdrop of magnetospheric conditions as described above, the more detailed sequence of events leading to the modelled substorm is now described. Most ground data used has five-minute spacing as supplied by Barbara Emery [personal communication, 1996]. In the five minute period from 1050 to 1055 UT on November 8, declines in the X component by about 500 nT were observed at College, Alaska, and at Dawson, Yukon Territory (see Figure 10.2; better time resolution is shown in Figure 10.9). Very close to 1050 UT, a dispersionless electron injection[2] was observed by the 1989-046 synchronous spacecraft, which was at local midnight (Figure 10.3). Dispersed electron signatures were observed later at two spacecraft at 8 hour local time spacings away from the injection site, and by 1984-129 10.5 hours of local time east (data not shown). These signals indicate that the onset took place near Alaska at 1050 UT, a finding consistent with the small amount of high time resolution ground data examined. There is no indication of growth phase from the electron data at any spacecraft, nor of an injection of any size in the several hours preceding 1050 UT. In the traces of Figure 10.1 it may be seen that this onset, marked by an arrow at about day 8.45, is preceded by lower levels of activity than were typical of much of the storm period. This is consistent with what the electron data suggest. An AL trace at larger scale (not shown) shows almost identical behaviour to that of the X component at College, so undoubtedly that station dominates the 89-station index. Since College is also one of the ‘official’ 12-station AE stations, the AE (when eventually computed) will likely give a good indication of onset time for this substorm.

---

[2] energetic particle data are from the service at http://leadbelly.lanl.gov/lanl_ep_data/ep_request.html

Figure 10.2 Magnetic observations at College and at Dawson on November 8 1993. Electrojet appears to be centred at College and to be somewhat north of Dawson.

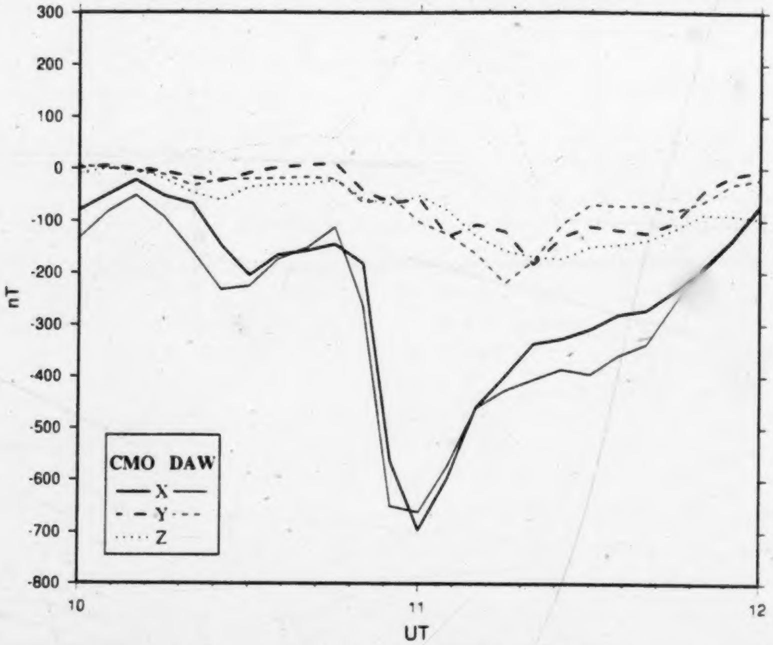


Figure 10.3 Dispersionless injection observed by SC 1989-046 on November 8, 1993. 1050 UT is marked. keV Channels (from bottom) are 50-75, 75-105, 105-150, 150-225, and 225-315. At highest energies the injection is preceded by a dropout.

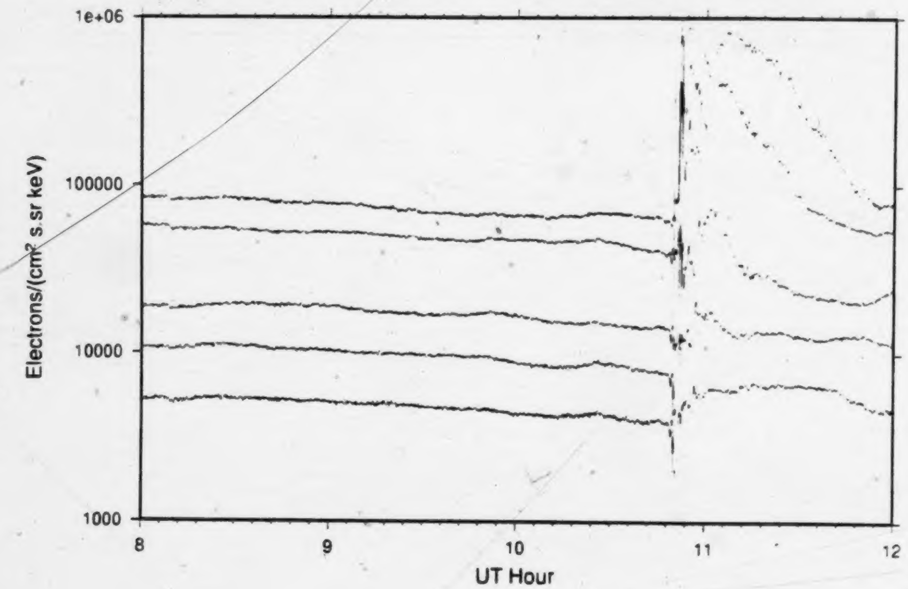
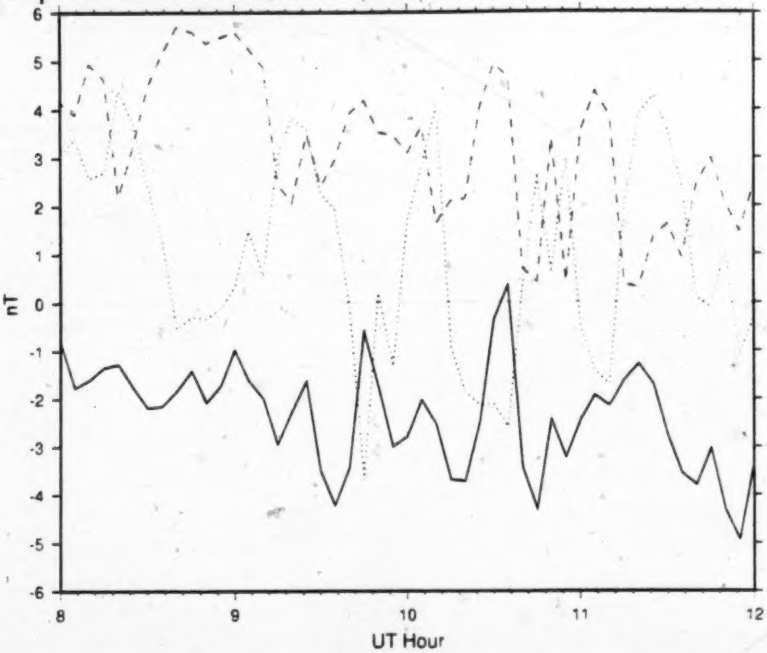


Figure 10.4 IMP-8 IMF in GSM coordinates from 0600 to 1200 UT November 8, 1993. Z component shown dotted, X solid, and Y dashed.



The generally northward IMF detected by IMP-8 and shown in Figure 10.4 is consistent with the relative quiescence of the electrojets before the 1050 UT onset. As mentioned above, no strong indications of previous onsets are present in the geosynchronous particle data either. This is in contrast to the latter half of the UT day, during which the IMF was southward a good deal of the time, resulting in several onsets being consistently indicated by the AL index and electron injection signatures. Unfortunately, data from the synchronous GOES spacecraft, located close to the region where the substorm current wedge would be expected, are not complete for this event, and will not be used.

There is some evidence in the AL index and in the electron signatures that the southward turning at 0620 to 0700 UT caused enhancement of activity (data not shown). The next significant southward turning took place from about 0940 to 1000 UT. This may have caused a small change in the AL index (with 89 stations, many of which are at this time near the midnight sector) but there was no electron signature. It seems safe to conclude that an expansive phase onset did not in fact occur. From 1010 UT to 1045 UT, the IMF had a relatively steady southward component of about 2 nT magnitude, and there was some indication of westward electrojet activity (perturbations of about 200 nT at College and Dawson) peaking at about 1030. Again no expansive phase onset appears to have taken place, although the enhanced activity may be classified as a pseudobreakup [Rostoker, 1997] as will be discussed below. It is the northward turning at 1040 UT that appears to have triggered the onset under study. At this time the IMP-8 spacecraft was about 20 R_E from Earth on the dawn side and only about 2 R_E tailward, so that propagation delays in the solar wind should not be a major consideration. The onset, already described, appears to have been isolated, with little structure in the ground magnetics (at 5 minute resolution) or in the electron signatures. Activity associated with the onset diminished at 1200 UT. Table 10.1 summarizes the fairly simple events in the 1000 to 1200 UT period.

Table 10.1 1050 UT Onset on November 8 1993

Time (UT)	Activity	Observational Basis
10	IMF, northward ~3 hours, returns northward after brief southward turning: no onset	IMP-8, AL, midnight sector stations, synchronous orbit electron detectors
1010	southward turning (2 nT) driven system activity	IMP-8 midnight sector stations
1010-1030	pseudobreakup signature	CMO, DAWS magnetograms
1040	northward turning X component decline at midnight	IMP-8 midnight sector stations
1050-1055	onset magnetic signatures dispersionless injection	CMO, BRW magnetograms SC 1989-097
1055-1200	negative X bay	AL and midnight sector stations

c. Modelling results

Data from 154 stations were assembled by Barbara Emery as part of the GEM Space Weather project and communicated to the author as acknowledged above. Two stations were found not to have useful data during the 1000-1200 UT November period of this detailed study, leaving 152 stations, of which most were in the northern hemisphere. The positions of these stations are shown in Figure 10.5.

The quiet day used in data preparation was 13 UT November 2 to 13 UT November 3. Although the D_{st} initial phase had already begun by the end of this period, it had not reached very large values (refer back to Figure 2.3). Toward the end of the two hour period considered here, the rise in D_{st} might lead to a bias in the D_{st}-like parameter used in modelling, and a possible inferred decline in D_{st} after the substorm may represent more the D_{st} positive phase biasing the quiet day than any real ring current effect. Quiet day removal was done by Barbara Emery, and data preparation after obtaining the data from her was minimal. Rotation into the centred dipole system was performed and some small problems with the data were addressed. These comprised a missing Z component (and both other components before 1030 UT) at Fort Yukon (FYU) and missing data values at the Greenland station Daneborg (DNB) at 1050 UT. To overcome the first problem, the missing Z values were replaced by the value -0 and models were run with an error value of 9999 on the Z component at this station to effectively remove all weight from it. Normal error values were used for the other two components since they contain useful and apparently correct data values from near the current wedge. Since the station contains important information in the two valid components, modelling was not started until 1030 UT when these became available. The second problem, that of missing data values at

DNB, was addressed by simply interpolating the values from 1045 and 1055. This station did not have significant perturbation values in any case. Data were transformed from all stations received, but only those from the northern hemisphere were included in the modelling. This left 121 stations, distributed as shown in Figure 10.5, which also shows the conjugate stations. In the figure, black dots indicate stations in the northern hemisphere. while larger grey dots are conjugate positions. The grid is 10° in latitude by 15° in longitude in the centred dipole system. Many stations shown have their exact positions listed in Table 6.8. Along the 30° line of centred dipole latitude, certain longitudes have been labeled. These indicate centred dipole longitudes for which meridian profiles may usefully be made.

Inspection of the data showed that traces of the evening sector eastward electrojet were weak. This situation appears to typify much of the GEM storm period as the small AU index values (see Figure 10.1) suggest. The morning sector westward electrojet was typified by X component perturbations of roughly -100 nT in North America at about 0300 MLT and 65°, while further east there is not much trace of it at this latitude. This is verified by inspection of data (not shown) from the Halley Bay conjugate station, which shows positive perturbations throughout the period. At higher latitudes (75°) there may be electrojet perturbations at roughly 0600 MLT but these do not extend appreciably further east as they are not prominent in the East Greenland chain which spans this latitude.

In the earliest modelling, a rough system of two electrojets and one substorm current wedge was made up and optimized with AFM. In these runs the evening sector electrojet did not settle into a realistic pattern due to the ambiguity of the data in that region. As inferred above from inspection of AU traces, this eastward electrojet appeared to be of minimal strength throughout the storm period. In such a case, the routine will move a weak system into another region where its parameters can be put to better use in reducing the chi-square. Little information is then gained about the global electrojet pattern. With the conclusion that the eastward electrojet was at many times too weak to be convincingly modelled, it was removed and only the two components with visible signatures were used. For times when the presence of a strong evening sector electrojet was indicated, it was put in, and this will be discussed in what follows. Initial conditions were set up, based on the earlier run, to be representative of conditions at substorm maximum. The initial current systems are shown in Figure 10.6. The current in the substorm current wedge was 0.75 MA and that in the westward electrojet 0.20 MA in this set of initial conditions.

As seen in Chapter 6, there are two ways for the routine to reduce the magnitude of perturbations associated with currents. One is to reduce the currents while retaining a configuration similar to that of the Hall electrojets. The other is to change the configuration to that of a north-south quasi-solenoidal system while retaining a large current. Here a constraint was placed on the system by penalizing the chi-squared if the current paths deviated much from being east-west. This method was found effective in eliminating the tendency to use north-south current systems to 'remove' excess current.

Figure 10.5 Locations of stations used in GEM modelling. Nadir 285.70. Insets show codes of stations in Fennoscandia and in Greenland.

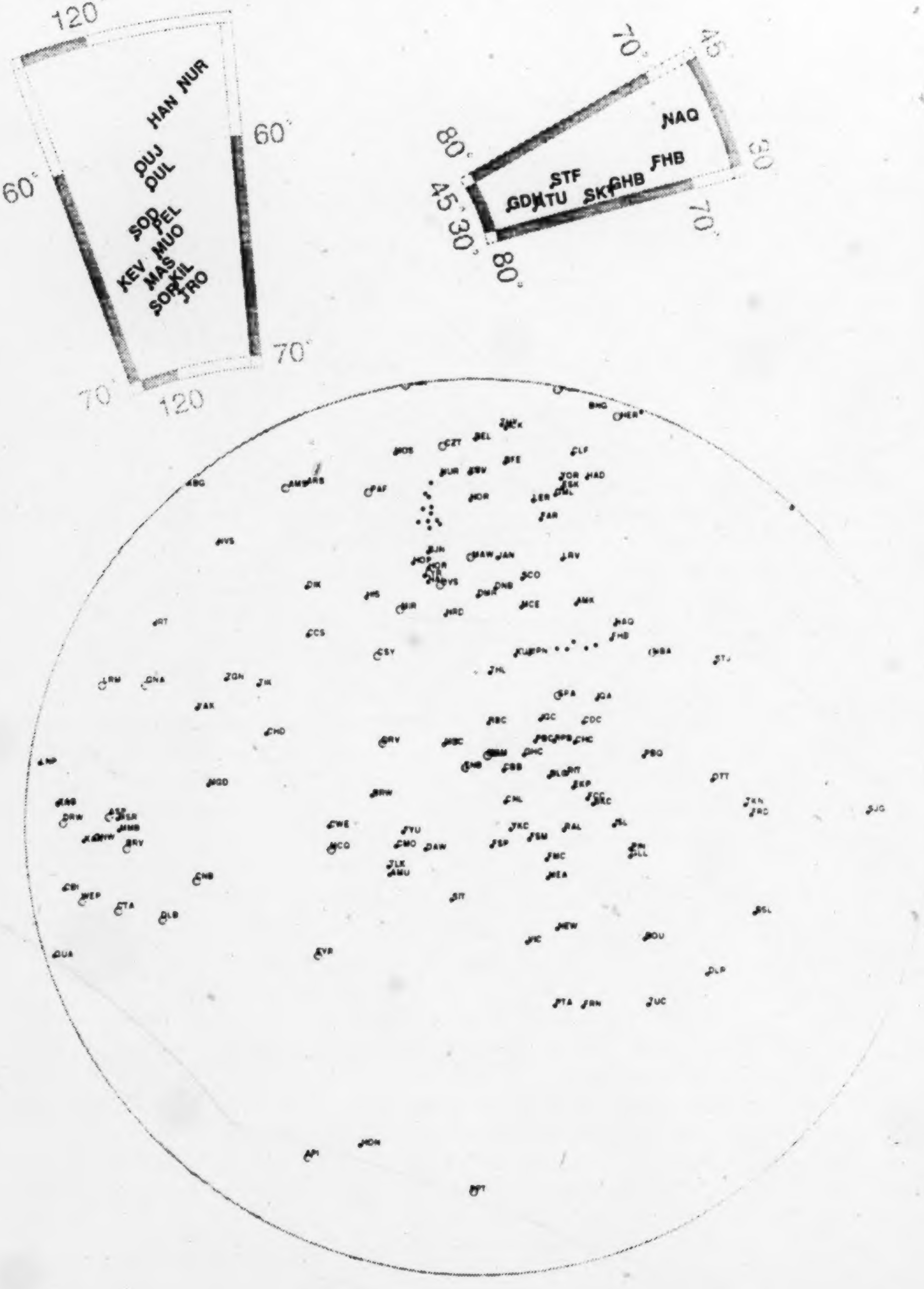

120°
60°
60°
70°
70°
120
HAN NUR
OUJ
OUL
SOD
PEL
KEV
MUO
MAS
KIL
SOR
JRO
70°
45
80°
30
45
30°
80°
70°
NAQ
FHB
STF
GDH
ATU
SKT
GHB

Figure 10.6 Initial conditions for modelling for November 8 1993.

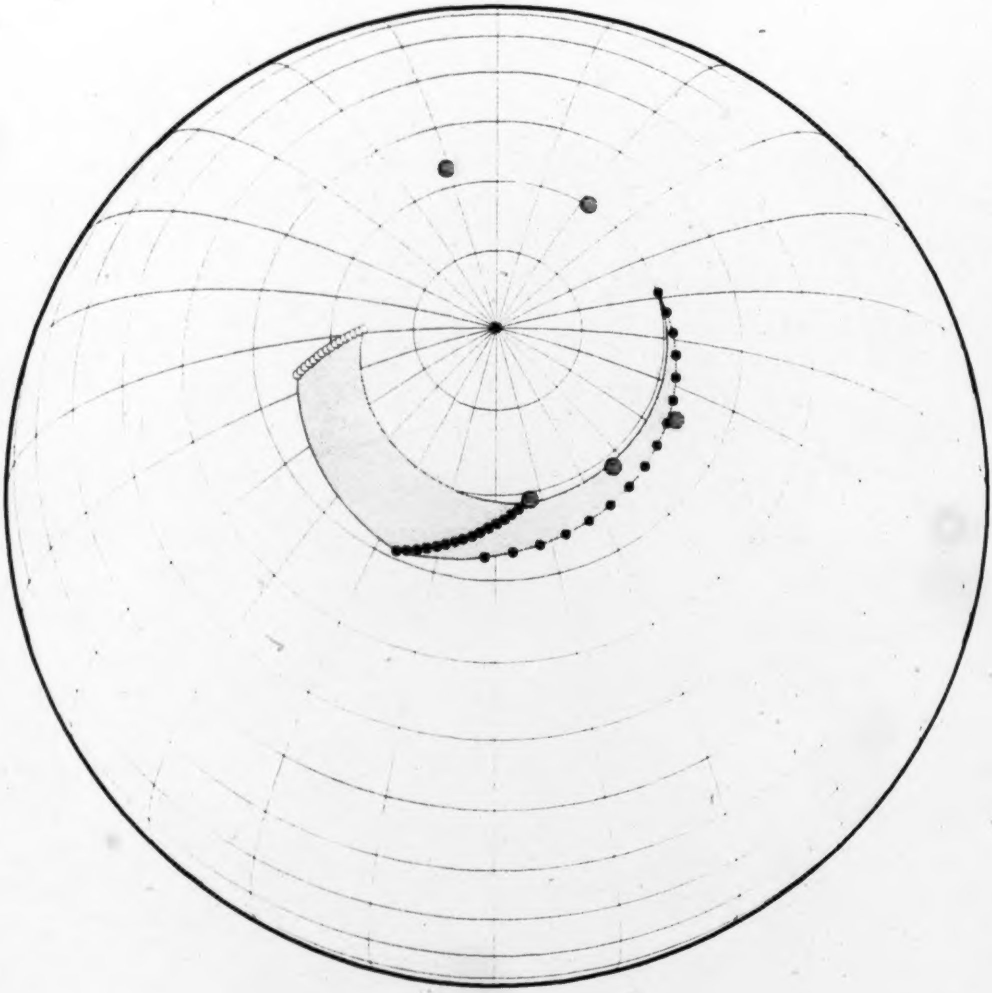
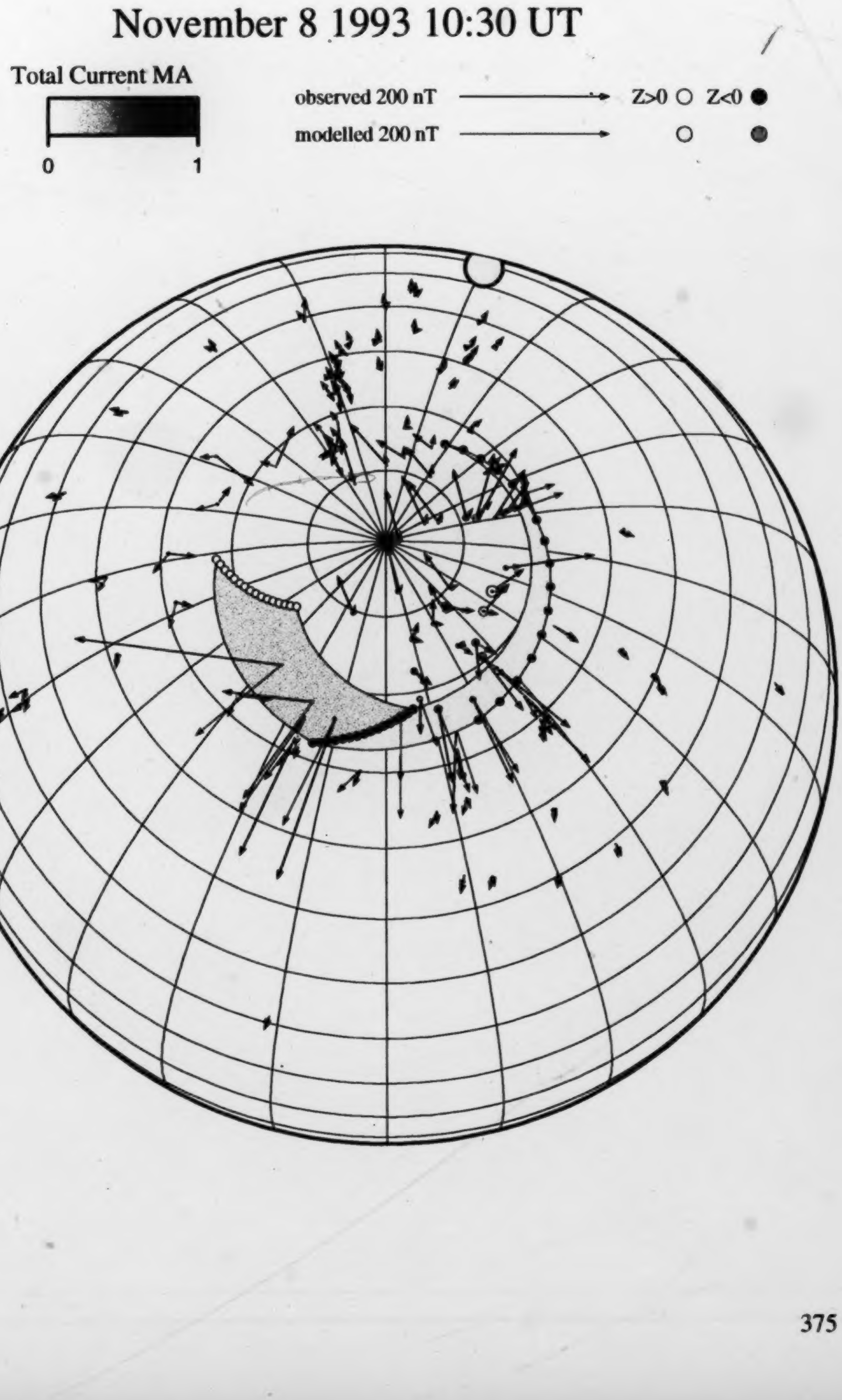
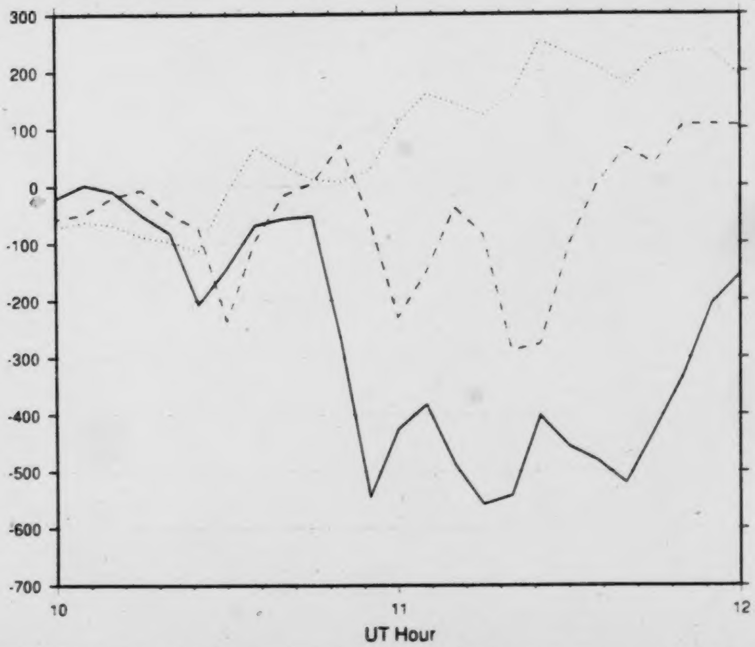


Figure 10.7 Modelling results at 1030 UT November 8 1993.


November 8 1993 10:30 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

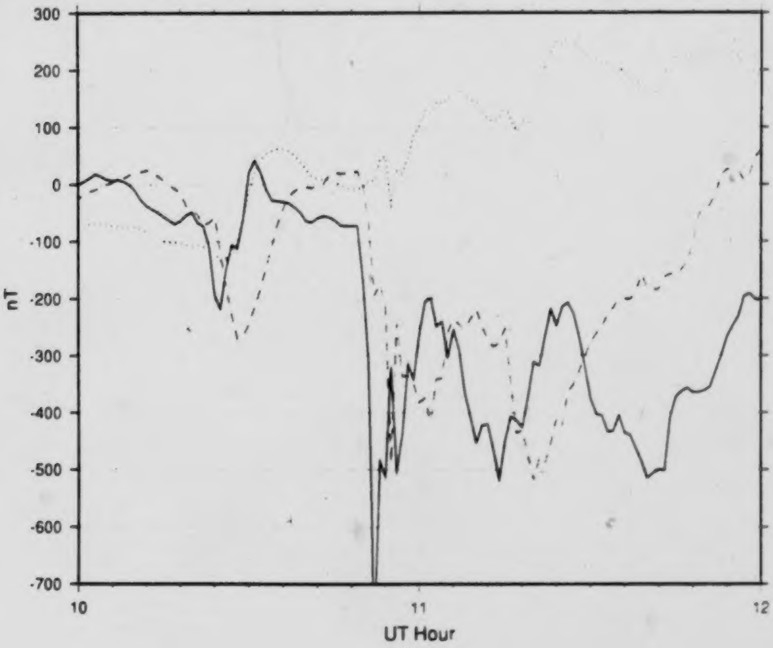
The results of modelling prior to onset are shown in Figure 10.7. In general, the agreement between model and observations is good. The polar cap perturbations, not well modelled, have a consistent duskward component while pointing in a generally sunward direction. These perturbations may be compared to those shown (as equivalent currents) by Friis-Christensen [1983]. The duskward tilt is consistent with those results for IMF $B_Y>0$ and B_Z zero or less than zero, which are the conditions in force at this time (see Figure 8.4, noting that B_Y is strongly positive, making the overall in-ecliptic component of the IMF basically duskward at this time). However, Friis-Christensen's statistical study indicates that under such conditions a strong evening sector eastward electrojet develops at high latitudes. This is definitely not the case here despite adequate station coverage. The morning sector westward electrojet is also not strong but is well modelled in the early morning part of the auroral oval. The 'current wedge' has had its current reduced to 0.24 MA from the initial 0.75 MA and the southward component has been generally well modelled under it. Although onset has not yet taken place, it is clear that the most westerly, and northwestwardly inclined, portion of the westward electrojet, already carries a stronger current than does the more easterly portion of the electrojet. The time point shown corresponds to the period of 'pseudobreakup' and strongest pre-onset current flow in this region.

Figure 10.8 Five minute averaged data from Barrow on November 8, 1993. Centred dipole magnetic components are shown: X_M solid, Y_M dashed, Z dotted.



It is also interesting to note the strong westward perturbations at Barrow and Fort Yukon, as shown in Figure 10.7. Such perturbations have been remarked on for other events studied in this project. The small separation of Fort Yukon from College, which does not show these westward perturbations, indicates that they are found within a relatively narrow latitudinal range. It should be noted that they are generally found in northwestern North America in studies shown in various chapters here, but that they are at differing stations in each case. In all of these cases North America was in the evening sector. The westward trend of perturbations would thus appear to be a real feature of the pre-onset evening sector. Although, as noted above, an onset does not appear to have taken place before 1050, currents in the Alaskan sector did strengthen around 1030. This may be seen by the slight decline in the X component at College and Dawson as seen in Figure 10.2, where the absence of a corresponding Y perturbation may also be noted. The perturbations at Barrow are shown in Figure 10.8, in centred dipole magnetic coordinates and with five minute spacing as used in modelling in this chapter. Comparison to Figure 10.2 will show the same decline in the northward component seen at those stations at roughly 1030, but also a pronounced westward component deviation. To illustrate that this is not an artifact of station data rotation, Figure 10.9 is also supplied, showing approximate perturbations (the zero level was adjusted to match that in Figure 10.8: these are not properly baselined) in XY geographic coordinates, with one minute averaging. It is seen that the perturbation in Y is strong and negative. Its centre appears to coincide with a reversal of polarity of the X perturbation and a level shift (upward) in Z. These are not the signatures of a simple electrojet enhancement, nor typically of a substorm onset, so it appears that some other sort of current system was present in the high latitude evening sector near the time of IMF southward turning and preceding a major onset.

Figure 10.9 One minute data from Barrow on November 8, 1993. These data are relative to the geographic system of axes.



This current system causing the X component, and possibly also the Y component, perturbations may be that of a pseudobreakup [Rostoker, 1997], with ground signatures which are very similar to those of an expansive phase onset. Inspection of Figure 10.3 will show, however, that there was no particle injection associated with this pre-onset event, while a very pronounced one did occur at the 1050 onset. In the one-minute data, the 1050 onset is very sudden and very sharp, and the temporal association with the injection is strong, while the possible pseudobreakup perturbations develop more slowly. Figures 10.8 and 10.9 also underscore an obvious difference between five minute-averaged and one minute-averaged data: the former reduce peaks and introduce temporal smear.

Proceeding with the global modelling, a more quiescent pre-onset condition is seen in Figure 10.10. The 'wedge region' westward current over Alaska has declined to .12 MA while the westward electrojet carries .096 MA of current. The Alaskan perturbations seem in general to be smaller than those indicated by the model although there is some variation. It is possible that a current close to .1 MA typifies the entire midnight sector electrojet. There may also be some indication of a weak evening sector eastward electrojet from northward perturbations at some stations there. As mentioned above, attempts to model this were not very successful: there is not enough 'power' in these perturbations to 'hold down' a model eastward electrojet and its parameters may be used better toward reducing the chi-square by placing it elsewhere. The Alaskan region also no longer shows strong westward perturbations in its more poleward stations: Barrow still has a significant westward component but the overall perturbation is smaller (halved from the 1030 observation), and Fort Yukon has near-zero perturbations now as compared to non-negligible perturbations observed nearby at College. The polar cap maintains its pattern of perturbations pointing toward the afternoon sector. These do not appear to have been associated with the pseudobreakup, which has by now declined in strength.

Figure 10.11 shows conditions right at onset time of 1050 UT, although the onset effect has been smeared out temporally by the five-minute averaging of the data used. The onset region reserved as the 'substorm current wedge' may now properly be called that, and in the model it carries .26 MA of current, substantially more than before. The westward electrojet has been reduced to only .054 MA, and southward perturbations near its intersection with the wedge are largely effects of field-aligned current. The central part of the westward electrojet is clearly at somewhat low latitude in the model, compared to where the large southward perturbations in northeastern Canada suggest it should be. A possible narrowing of the westward electrojet is not obvious. The apparent narrowing since 1030 in the Alaskan sector may be spurious since the northern border of the electrojet is not well constrained by stations. Again, there are signs of a weak eastward electrojet south of the substorm current wedge. The polar cap perturbations appear to have decreased in magnitude and veered sunward with the onset: these are not represented in this model which does not include Region 1/2 currents. Finally, subauroral perturbations are matched in direction but not in magnitude at Sitka near the downward wedge current (where the magnitude is too small) and at Magadan equatorward of the upward wedge current (where it is considerably too large). At most North American stations midlatitude perturbations are well matched.

Figure 10.10 Modelling results at 1035 UT November 8 1993.

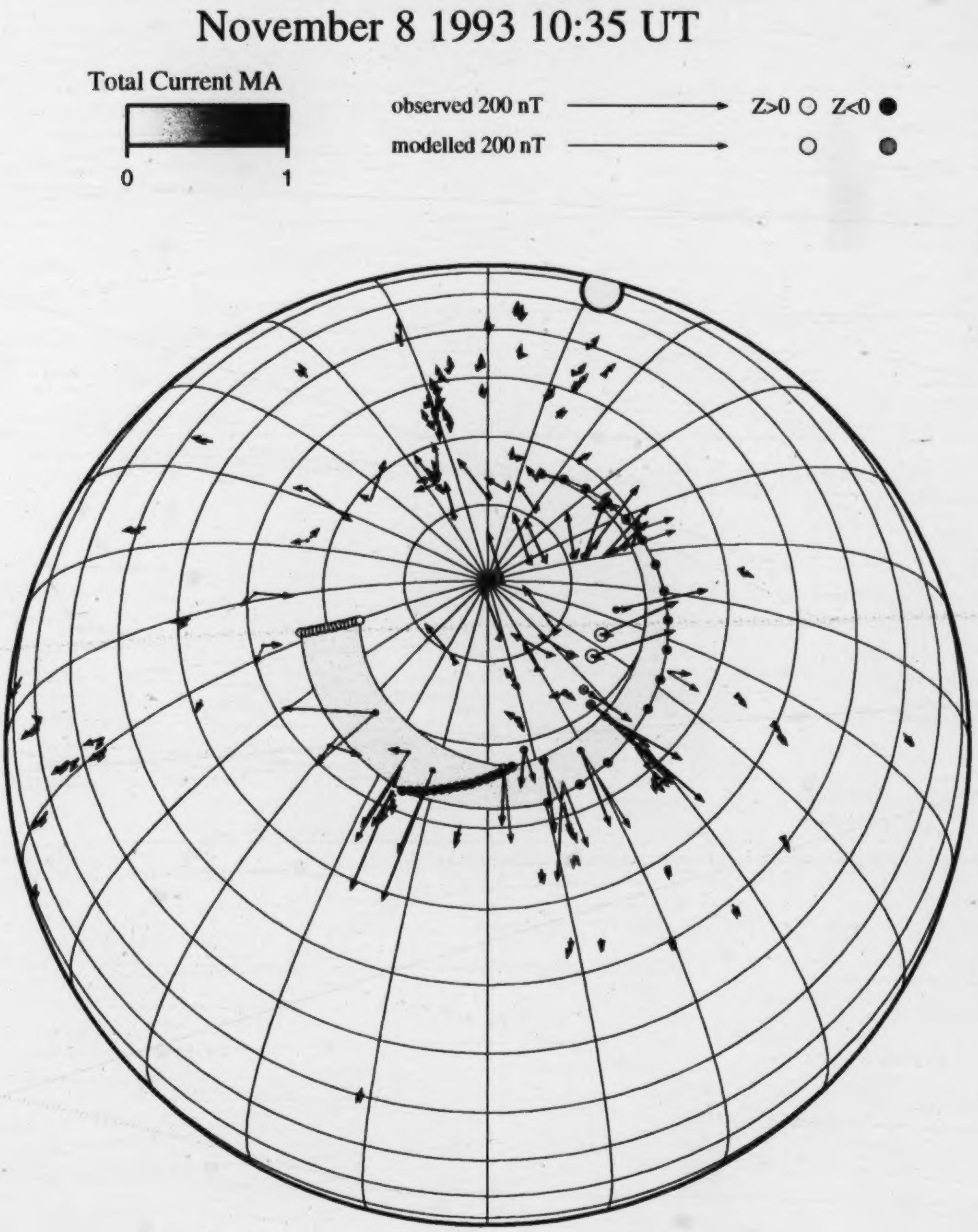

November 8 1993 10:35 UT
Total Current MA
0
1
observed 200 nT
Z>0 ○ Z<0 ●
modelled 200 nT
○ ●

Figure 10.11 Modelling results at 1050 UT November 8 1993.

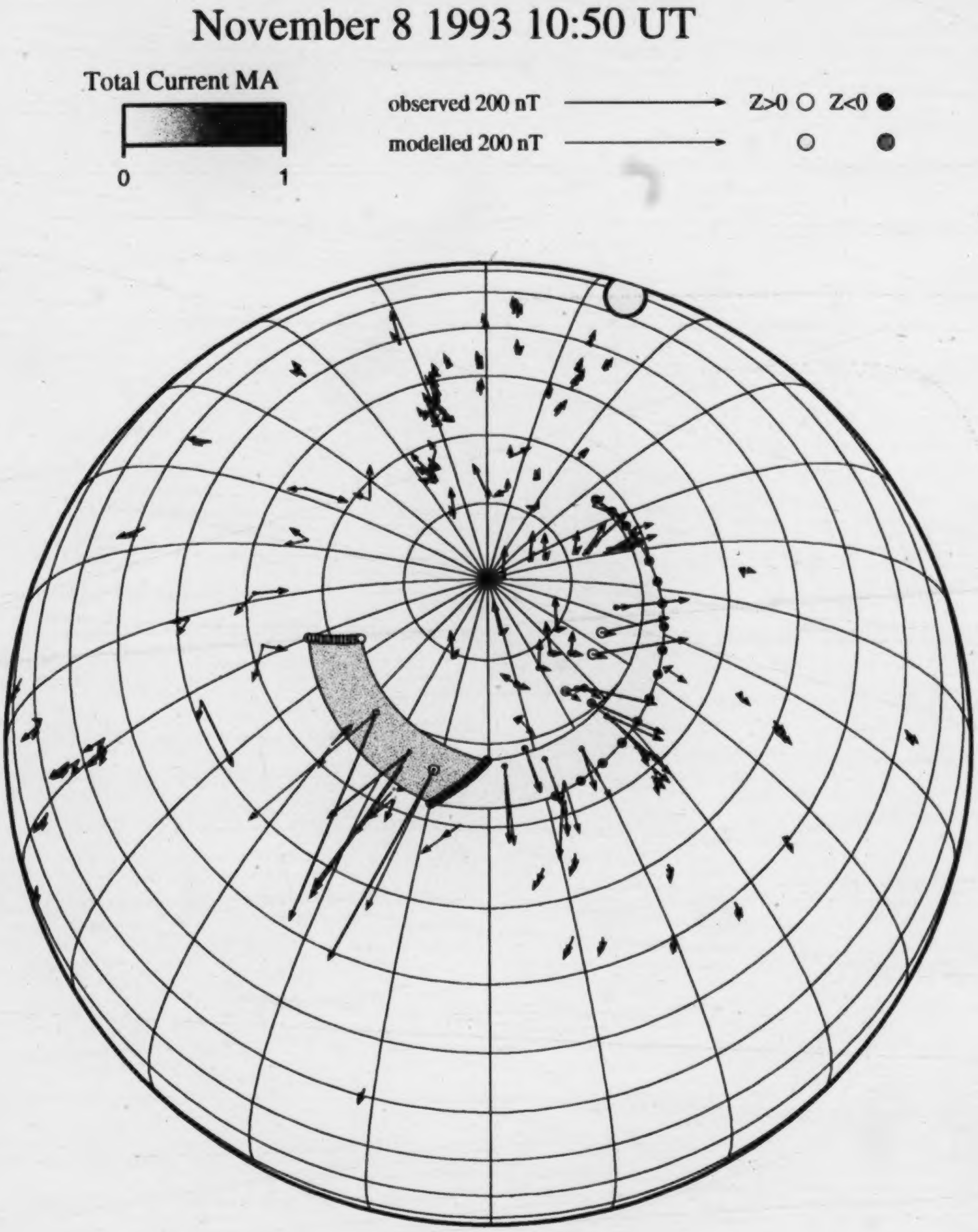

November 8 1993 10:50 UT
Total Current MA
0
1
observed 200 nT
Z>0 ○ Z<0 ●
modelled 200 nT
○ ●

Figure 10.12 Modelling results at 1055 UT November 8 1993. Strong wedge obvious: perturbations due to eastward electrojet much enhanced.

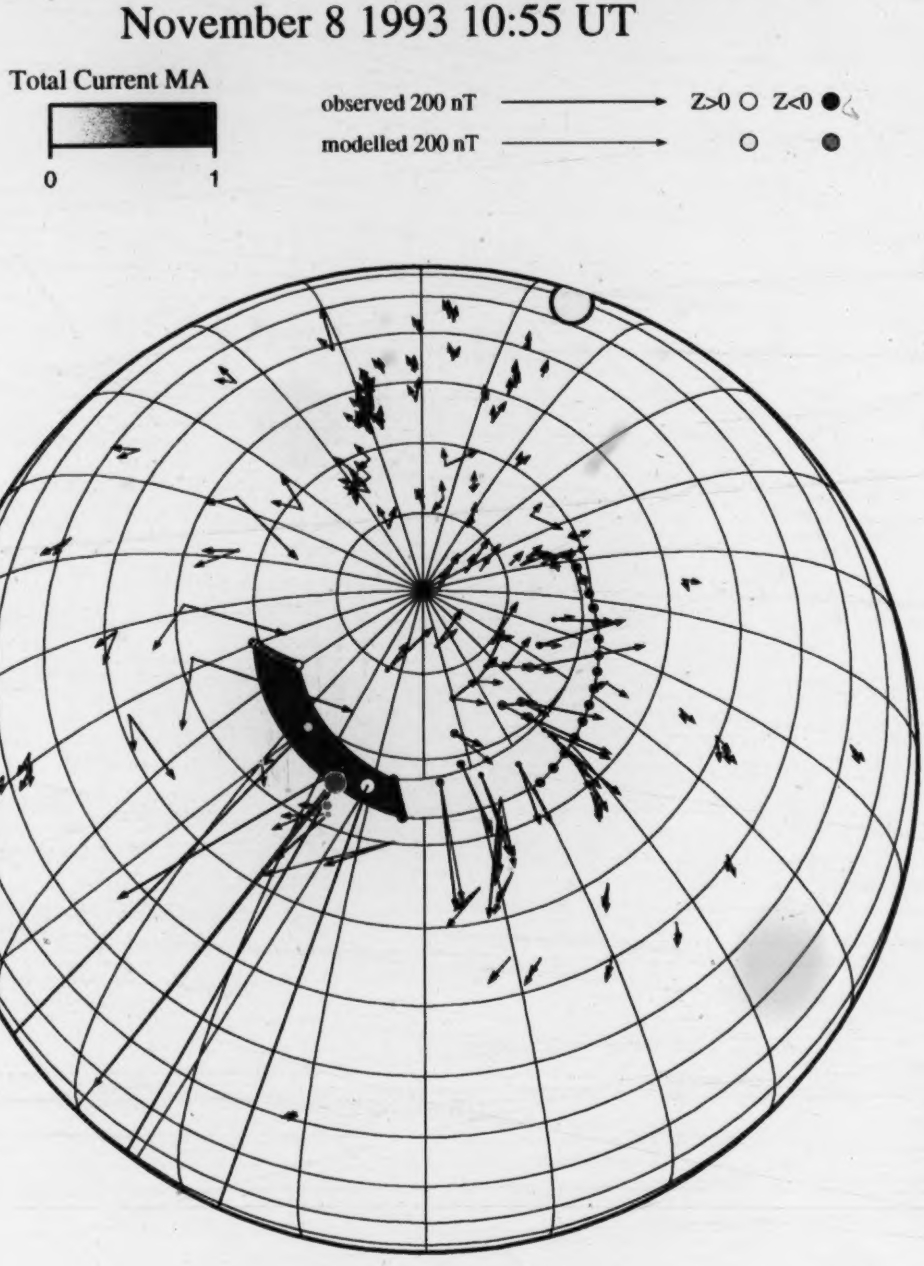
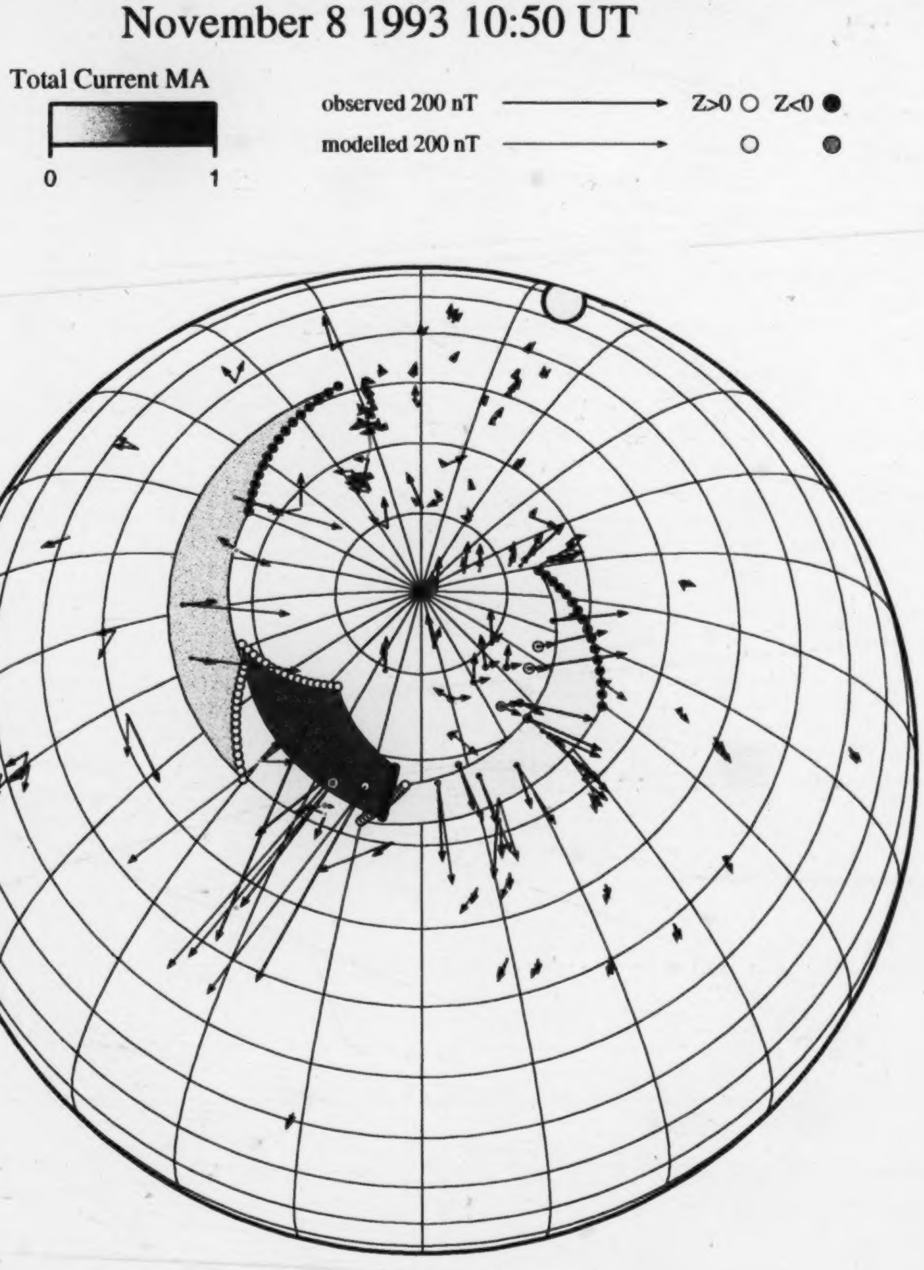

November 8 1993 10:55 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Figure 10.12 shows the situation at 1055. With higher time resolution data it is likely that the situation about one minute (rather than five) after onset would be close to this. The general features from the previous figure still apply, but are more pronounced: enhanced current in the wedge system, good match in the western part of the westward electrojet while the eastern part should be further north, veering of polar cap perturbations sunward, and subauroral perturbations associated with the wedge field-aligned currents. Some model and perturbation arrows go outside the plotting region but these large perturbations are in fact well matched. For example, the observed and modelled X_M results at Barrow are -544 and -608 nT, while those at College are -564 and -455. At Fort Yukon, however, the match is less impressive, with -254 nT observed and -647 nT modelled (accounting for nearly 10% of the weighted error at all stations). The most important change from the 1050 UT model is that the perturbations clearly associated with the evening sector eastward electrojet are much enhanced. One might expect that these could now be modelled and this is explored below. In quantitative terms, the wedge current is now found to be .63 MA, while that in the westward electrojet has decreased yet more, to .036 MA. Once again, it seems that the latter result is somewhat spurious in that the pattern of perturbations in western North America (including at subauroral latitudes) appears to be dominated by the downward field-aligned current associated with the wedge, while the electrojet is likely at higher latitudes in eastern North America. That more easterly part of the electrojet likely carries a larger current than the model indicates.

The model done for 1100 UT resembles that at 1055 to a large degree and is not shown. The wedge current at that time had increased to .74 MA. The recovery phase of this isolated substorm was not modelled.

From the foregoing, it is clear that the eastward electrojet in the evening sector essentially turned on at or shortly after the time of onset of this substorm. Despite, or to circumvent, the failure of the initial models to properly use the parameters of an eastward electrojet from a rather general initial condition, another modelling run was attempted with an initial condition close to that which would be inferred by inspecting the foregoing figures. It was guessed that the current in the system was .3 MA and that it was roughly over the stations in the evening sector which show northward perturbations. All parameters were allowed free variation in these models, whereas the previous ones had the westward electrojet and wedge constrained to have a common line of intersection. The results of this modelling are shown in the next two figures. Figure 8.13, for the 1050 onset time, shows the freer variation taken advantage of: the junction of westward electrojet and wedge current is not exact, allowing a bit more geometric leeway to each and a better match to observation. As discussed previously, the actual physical analogue of the model in this situation is a feeding in of current to the wedge to supplement that of the electrojet, and it is unlikely that there is actually upward current flowing in the region. The net downward current is the difference between the upward and downward currents in the model. The new wedge current is .50 MA and electrojet current .061 MA. The former is larger than the previous result of .26 MA since the current is spread out more as the wedge now extends further northward, while the latter is comparable to the previous result of .054 MA. With few stations under it, the eastward electrojet does not appear to have optimized properly.

Figure 10.13 Modelling results including eastward electrojet at 1050 UT November 8 1993. Current in the eastward electrojet appears too large.



November 8 1993 10:50 UT

Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

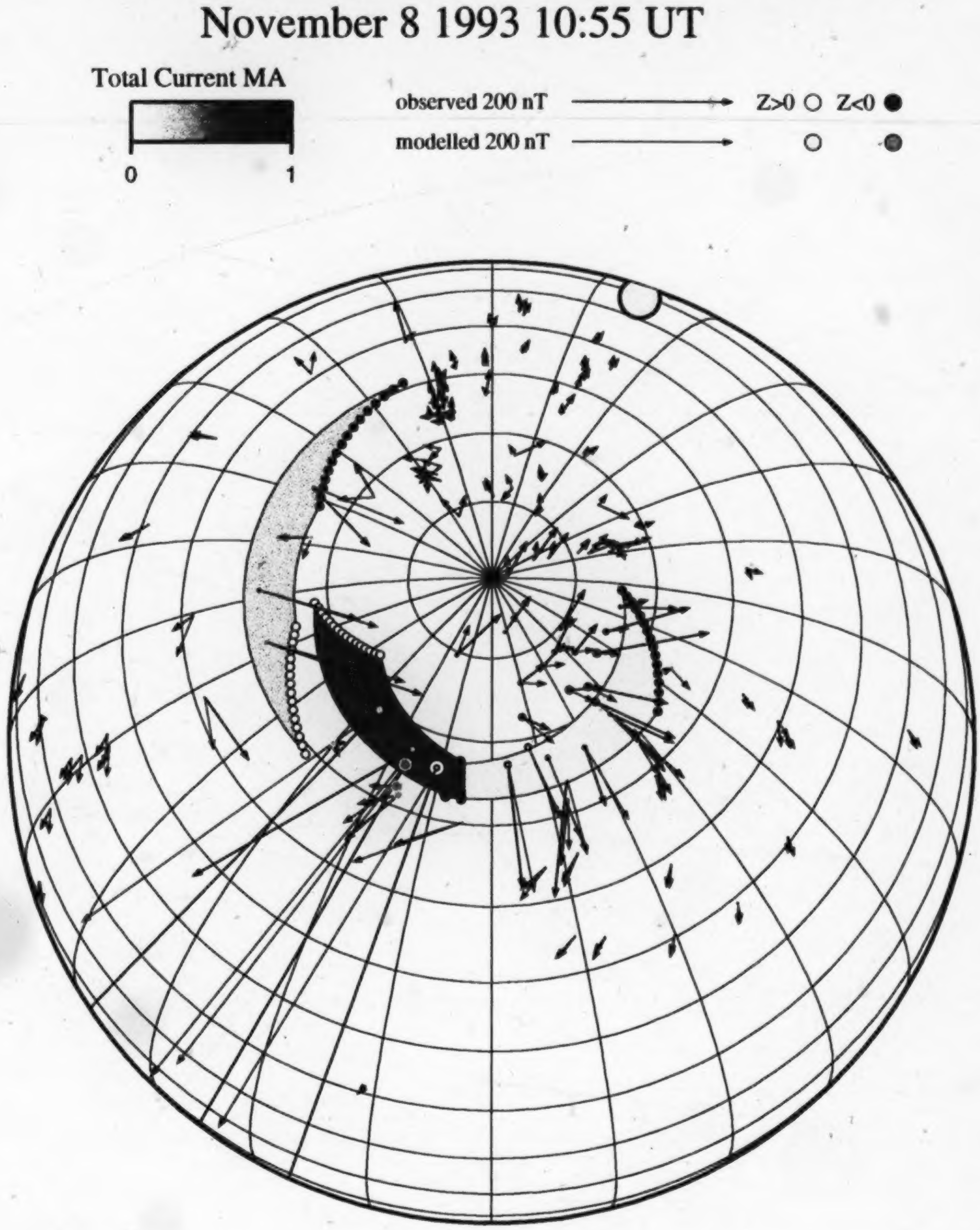
Starting from a rather large current initial condition, the evening sector eastward electrojet clearly has a current too large to account for the perturbations at this time. It has been reduced only to .23 MA from an initial .3 MA. The likely explanation for not having a further reduction is that the subauroral station at Magadan has a weighting which overall is about twice that of all the auroral zone stations in the evening sector. This is since the modelling rules attempt to balance small size of the perturbations seen at midlatitudes by assigning a higher weight. If such a station shows an anomalously large perturbation, as is the case here, the results can be quite skewed by it.

Figure 10.14, on the other hand, shows excellent agreement between the evening sector perturbations observed and modelled. Clearly the eastward electrojet did strengthen, and the modelling results show that .24 MA of current were flowing in it. The perturbation at Magadan has been relatively well matched and thus does not appear to have had a large negative influence. The wedge current dominates, but at .57 MA seems slightly underestimated. This is shown by the fact that the Alaskan stations have perturbations about 20% too small in the X_M component. The morning sector electrojet appears to model most perturbations near it well, with its eastern end likely needing to be further north.

Models which include the three standard current systems of morning and evening sector electrojet and wedge have been presented. Station coverage in this modelling is good, although there exist gaps over oceans. Reference to the recent figures shows that the dayside does not have much current flowing on it as indicated by the model. This has been a common aspect of most of the modelling performed to date. For example, it is a feature of the major CDAW 9B event of Chapter 6. In other events, generally selected to have data on the nightside, it has been difficult to justify the conclusion that not much current was flowing on the dayside. Here, there is good global coverage. The dayside data are present yet still do not indicate currents of any appreciable strength flowing there. The overall view of the auroral oval presented, without current flow on the dayside to any appreciable extent, appears to be correct, and in this case supported by the magnetic data.

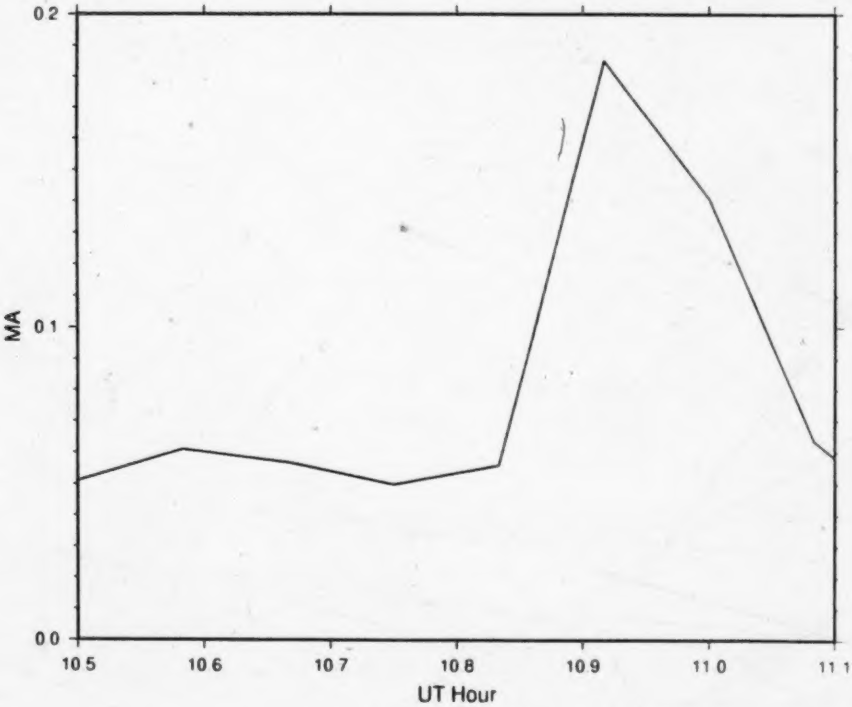
In the modelling, A D_{st}-like parameter was used as has always been the case in global modelling in this thesis. In the two current system runs from which the most results have been presented, this parameter varied from about -22 to about -27 nT between 1030 UT and 1100 UT. Barbara Emery [personal communication, 1995] has provided a file of 27-station D_{st} which varied from -29 to -34 nT during this same time period, and the WDC-C2 provisional D_{st} from the four stations SJG, HON, KAK, and HER, which remained very close to -27 nT. The behaviour of the D_{st} in the program resembles that of Emery's low latitude stations, while being of slightly lesser magnitude.

Figure 10.14 Modelling results including eastward electrojet at 1055 UT November 8 1993. Perturbations due to eastward electrojet are well matched.


November 8 1993 10:55 UT
Total Current MA
0
1
observed 200 nT
modelled 200 nT
Z>0 ○
Z<0 ●

Since the evening sector eastward electrojet appeared to respond to the onset, as seen in the preceding figures showing perturbation vectors, and since its weakness before onset precluded modelling it well, it was decided to attempt to model only the eastward electrojet using the stations which showed a clear signature from it. These stations were DIK, TIK, and CHD. The further assumption to allow modelling was that the electrojet did not move: then the only parameter to be solved for was the electrojet current. The electrojet configuration was held fixed with the geometric parameters from Figure 10.14 used. The resulting variation of its current is shown in Figure 10.15. This quantitatively shows the development of the electrojet current and is in agreement with what may be inferred from the perturbation arrows shown in the figures depicting modelling results. The current was nonzero but small prior to the onset, and rose from that level of about 0.05 MA to about 0.2 MA at its maximum, which was one time step (five minutes) after the expansive phase onset. The modelling has been taken slightly past 11.0 UT to show that the current after that time appeared to be in decline once more, apparently approaching pre-onset levels.

Figure 10.15 Total current in the eastward electrojet on November 8, 1993.



d. Summary

This event is the simplest studied in this thesis and appears to conform to the paradigm established by CDAW 6. Southward turning of the IMF allowed energy to be stored in the magnetosphere, which was unloaded at northward turning through a particle injection and enhanced currents in the substorm current wedge. As an isolated event, albeit during a storm, it does not seem to have involved formation of the double auroral oval. Any discrepancies between the position of the westward electrojet found in the model and the locations in eastern North America where its perturbations were observed can be explained by a bend in the electrojet rather than a doubling.

Other notable aspects of the event phenomenology include that:

1) the evening sector eastward electrojet was nearly absent prior to the 1050 onset, strengthened shortly after it, and declined once more

2) there were few indications of a growth phase in the geosynchronous orbit energetic electron data

3) about 20 minutes prior to onset there were strong westward perturbations under the northerly part of what would later be the substorm current wedge's ionospheric current flow region. These had many of the characteristics of what is described as a pseudobreakup, and should not be confused with high latitude onsets seen in other events

4) current flow rose abruptly from about .1 to over .7 MA in a current wedge of roughly 45° longitudinal span, at roughly 70° centred dipole latitude in the midnight sector but tipped toward the northwest

5) there was little response of the morning sector westward electrojet to the onset

6) the onset was clearly triggered by a northward turning of the IMF, with the southward period having lasted about 20 minutes.

During the event, good station coverage allowed the longitudinal extent of the electrojets to be firmly established and they were basically confined to the nightside.

Certain technical aspects of the modelling are also worth noting. The large number of stations available made the models slow to run compared to those in other chapters which generally had only about half the number of stations available. Many of these stations did not contain 'positive' information about the electrojets, although that lack allows the region where they were not found (the dayside) to be reliably established. The data's five minute spacing spreads out the perturbations temporally, so that the onset seemed to be spread out over a longer period than the one-minute data would suggest. The global routine was unresponsive to the weak pre-onset eastward electrojet. Regional modelling could be used at this time, and global modelling after it strengthened near onset time.

Chapter 11. Discussion and Conclusions

The work done here has both scientific and technical implications. The new AFM technique has contributed to the interpretation of magnetospheric behaviour, but its technical aspects need to be considered in interpreting the results. The first sections of this chapter will emphasize the results and the latter sections the use and promise of the technique.

As discussed in the introduction to Chapter 4, simple models lead to a parametrization of substorm disturbances as arising from a 'wedge' system of currents in the magnetosphere and ionosphere. This term is usually attributed to Clauer and McPherron [1974], although those authors are careful to point out that their paper reflects the concurring views of many researchers of the time. In this view, the configuration of the wedge, while allowing a convenient parametrization, also has physical significance in that the flow of currents is attributed to disruption of normal current flow in the (possibly near-Earth) magnetotail into the ionosphere via field-aligned currents. Although there are, as detailed in the preceding chapters, many manifestations of substorms, and as we have seen, substorm effects on all of the auroral zone current systems, the current wedge is the system most closely linked, at least in researchers' minds, to the substorm. For that reason, it merits emphasis in this chapter.

a. Characteristics of the Substorm Current Wedge

As modelled in this work, the substorm current systems can be complex, consisting of at least the current wedge and two 'driven system' electrojets. It has also been shown that 'onset-like' activity can occur at high latitude and away from the midnight sector and not necessarily be reflected in the AE index with its known biases [Baumjohann, 1986]. The generally-accepted view is that there are indeed two current systems, wedge and driven, and that these reflect different aspects ('physics') of the interaction of the solar wind with near-Earth space. In this view, driven system activity can mask wedge activity if sufficiently large, at least as viewed by the AE index [Fairfield, 1991], possibly giving an erroneous view as to the nature of substorm onsets. It has been illustrated here that a similar problem also occurs. Even expansive phase activity of larger magnitude than associated with the driven system may not be reflected in the index if its relatively localized zone of occurrence is far from AE stations, an effect exacerbated by the fact that the Z component does not figure in the AE index at all. The expansive phase activity generally accepted to occur and seemingly most studied in recent years is generally at a latitude which maps to the relatively near-Earth magnetotail and is well reflected in the AE index as a consequence [see again Fairfield, 1991]. Such activity may also convincingly be shown to be associated with 'unloading' of stored energy from the magnetotail. However, several cases have been demonstrated of activity associated with poleward border of a double auroral oval. In many respects this activity is reminiscent of substorm onsets, but there are biases against detection of many of the signs of onset associated with it. These biases are twofold. As mentioned above, the AE stations sample a relatively low latitude in terms of the double oval, and would primarily detect its southernmost portion. Further, events detected by AE likely map relatively near the Earth and may have their signatures detected by synchronous spacecraft. As a result, the onset-like nature of activity on the poleward border of the auroral oval, or on the more poleward of the two auroral ovals in the double auroral oval scenario, is not generally acknowledged. Here the use of comprehensive ground data has led to the finding that many of the characteristics of expansive phase onsets are found not only near the latitude of AE stations through the formation of the commonly accepted substorm current wedge. but also at higher latitudes which appear to generally correlate with the poleward part of the double auroral oval. This type of onset was noted in Chapter 6 both before and after major onsets which took place in the southern auroral oval. There, it also appeared that an onset took place at the poleward border (in the morning sector) which was as large in intensity as those normally associated with the midnight sector main oval, and which produced very long-lasting (several hours) perturbations. During the major substorm sequence of Chapter 8, there was a minor amount of high latitude onset activity. In contrast, the relatively short and isolated events of Chapters 7, 9, and 10 did not show double ovals and thus did not have high latitude onsets on their poleward borders. In Chapters 9 and 10, there was precursory activity evident, but this seems to have been in the same general location as the incipient substorm current wedge in each case. Such activity falls under the category of pseudobreakups, and is not the same as the small onsets under discussion. The most important conclusion from finding that there are onset-like features on the poleward border is that the substorm expansive phase process can occur near either border of the electrojet.

While substorms reflected by the AE index likely map to the near-Earth magnetotail, northern border onsets likely have to do with the plasma sheet boundary layer (PSBL). Frank [1988] (using data of collaborator Craven) demonstrated, using ISEE 2 at relatively high latitude in the near-Earth tail (GSM X=-7.5, Y=-2.7, Z=-4.8) that the PSBL maps to the more poleward discrete arc set seen in simultaneous DE 1 imaging. These observations were made during an active yet relatively stable period, and the crossing of the plasma sheet boundary layer is very distinct, with large particle density decreases as the lobe is entered. Due to the nearness of the spacecraft to Earth, the mapping (with the Tsyganenko-Usmanov model; see reference in Frank [1988] or Tsyganenko [1997]) is likely reliable.

Much has been made of northern border onsets. It is also clear that activation at or near the southern border can occur. This has been shown here notably with meridian chain results in which the electrojet narrows at onset, generally into the equatorward part of its previous extent. The two activation regions identified support the idea of a 'double auroral oval' and suggest that activity in either the central plasma sheet or in the PSBL may lead to substorm onsets which are similar in many of their characteristics. High latitude substorms often do not possess some of the classic substorm signatures such as particle injections, but this may, at least some of the time, be considered to be a selection effect due to the usual detectors of such particles being located at synchronous orbit and therefore not properly located to detect them when they are associated with high latitude substorms. The view that the substorm current wedge reflects diversion of cross-tail current into the ionosphere is not necessarily affected by the latitudinally segregated nature of substorms; one must simply consider that there are two likely layers of the tail from which such diversion can occur. In seeking the physical mechanism behind substorm onset this duality, with similar behaviours arising in regions with quite different characteristics, will have to be considered.

The observation that there is a double auroral oval with high latitude substorm signatures at its poleward border also reflects upon the classical picture of equatorward expansion of the polar cap/auroral oval system in response to southward IMF. If there are two active regions, it is not possible in all cases to claim that the observation of substorms at high latitude and subsequently at lower latitude in fact suggests that the auroral oval expanded equatorward. One must carefully disentangle the position of the substorms considered, relative to the overall auroral oval, before making inferences about temporal behaviour of the entire oval.

b. Aspects of the Global Electrojets near Onset Time

Despite the fact that the pre-onset electrojets are likely associated with the driven system and thus activated through quasi-continuous input of solar wind energy into the magnetosphere, while the expansive phase substorm current wedge is associated with an unloading process for stored energy from the magnetotail, the dramatic growth of the wedge appears to affect the electrojets even in regions quite distant from it. These effects have two forms: the position and/or current strength in the electrojets may be affected. In the region where the substorm current wedge formed, there generally already was an electrojet signature present [de Groot-Hedlin and Rostoker, 1987]. It goes without saying that this signature is dramatically changed at onset time.

The use of AFM with a simple electrojet model indicates that significant latitudinal motion of electrojets can occur during active periods. These motions may be of two types: a bulk motion (generally of both electrojet borders) or an apparent motion likely due to the activation of a narrow current filament. The latter can give rise to an apparent narrowing of the electrojet also. The observed motions take place both in the pre-existing electrojet in the region where the current wedge forms, and in distant regions.

During the late growth phase of the CDAW 9B event of Chapter 6, southward motion of the eastward electrojet in the evening sector took place, with correlated changes in the electrojet profile and at synchronous orbit (as noted by Skone [1994]). Evening sector southward motion also took place in a strong evening sector electrojet prior to onset in the CDAW 9A event described in Chapter 7. The evening sector electrojet during the CDAW 9C event of Chapter 8 was not modelled by the more appropriate meridian profile inversion approach, but global modelling indicated little change in its latitude. There was no indication of southward motion of the rather weak eastward electrojet on August 8, 1991, but again the distribution of stations was not adequate to indicate whether it took place or not. For the remaining two events such motions were not investigated. The consensus from this small sample is that southward (i.e. equatorward) motion of the evening sector electrojet may take place if that electrojet is relatively strong.

If the substorm current wedge is considered as a uniform entity, the poleward motion of its northern border is a well-known phenomenon, and is confirmed by auroral imaging [Rostoker *et al.*, 1987]. Such motion has been observed here and will be discussed in more detail below. A more subtle, yet perhaps important, result is that the electrojet borders come closer together at onset time and possibly before what most would regard as onset time. This effect was noted in four of the six events studied. The exceptions were the CDAW 9A event and the GEM event, both of which were the simplest events studied, having a 'classic' development with southward turning of the IMF, short growth phase, and onset. Perhaps also of significance in this context is that these were the only events with clear solar wind triggering of the onset (of four events with IMF data available). It could be concluded that the narrowing of electrojet borders is evidence of internal triggering. Ohtani and coworkers [1992] examined substorm onsets as observed by AMPTE CCE in the equatorial plane and propose that an 'explosive growth phase' takes place about one minute before dipolarization. During this short phase, the crosstail current is strongly enhanced within a small radial region, after which an instability is triggered causing crosstail current disruption. The onset locations, observed *in situ* by the spacecraft, were within 9 R_E of the Earth. The spacecraft did not, however, sample space beyond this radius. Nevertheless it is interesting to note that this spacecraft study implies a radially confined event, relatively close to Earth, of short duration immediately preceding onset. A relationship to the necking observed in this study is possible. One must proceed with caution, however, since the electrojet in the cases without narrowing of the electrojet borders was relatively narrow to begin with. Since it has been noted that activity is possible in each border region of the double auroral oval, it may simply be that only one auroral oval is active in cases where there appears to be no electrojet narrowing. In the other case, the apparent narrowing would be due simply to the activation of one oval (the more equatorward) or of regions within it, without enhancement of activity in the poleward oval.

Another observation in regard to electrojet borders is that there may be equatorward motion of the equatorward boundary of the wedge electrojet, or of the westward electrojet east of the wedge region, at or shortly after the time of onset. Equatorward motion of this boundary took place for the same four events for which narrowing of the electrojets was observed, did not take place for the CDAW 9A substorm, and was not a significant feature of the global modelling solution from Chapter 10. Again there seems to be a connection to IMF triggering: in the two events which were clearly triggered by an IMF northward turning, no southward motion took place.

Another aspect of the electrojets cannot necessarily be considered to change near onset time. This is the angle of the evening sector westward electrojet, which in all cases was found to exist north of the eastward electrojet. This is a well known aspect of the configuration of the electrojets as previously studied by meridian chain studies [Rostoker and Phan, 1986] and by global modelling (see Kamide [1988]). Here it can be made quantitative in that the westward electrojet exists as a distinct entity in the global AFM modelling. In most cases, the evening sector westward electrojet has its latitude increase as it flows westward. The one exception is during the event of August 8, 1991, when the most pronounced characteristic of the eastward electrojet in the evening sector was its weakness. Rostoker and Phan (refer back to Figures 1.9 and 1.10) found a difference in the configuration of the westward evening sector electrojet which seemed to not veer northward as much during times of higher activity, presumably near the time of onsets. No such difference has been found from global modelling here: rather the current within the electrojet usually strengthens and continues to occupy the same approximate region as it had previously. In some cases it has additionally been found that the current in the eastward and westward electrojets also strengthens at or after onset.

Another change associated with onsets and the electrojets is that in the Region 1/2 system, which is known to be spatially colocated with the electrojets [Senior *et al.*, 1982]. The short time scale variation within this system is studied with difficulty since ground signatures are not generally recognized and since spacecraft, the best indicators of the Region 1/2 system, traverse it sporadically. In several cases here, it has been found that the polar cap sunward perturbations can be most consistently explained as due to the Region 1/2 currents. For the CDAW 9B event of Chapter 6, these perturbations did not appear to change dramatically at onset time, and the overall configuration of currents responsible for them is primarily that of the driven system electrojets. DMSP showed a strengthening of the Region 1/2 currents in the midnight sector, and concentration at the northern border. However, the lack of change in the polar cap suggests that Region 1/2 changes were not dramatic on a global scale. For the discreet and pronounced onset of CDAW 9A studied in Chapter 7, global modelling including the field-aligned currents was not done for a time sequence of models. It was shown only that addition of the Region 1/2 systems improved the polar cap fit after the time of onset. Independently, however, DMSP magnetic data were used to suggest that the early morning sector Region 1/2 currents turned on (i.e. greatly increased their total current) at onset time. In the evening sector, however, ground modelling to match perturbations in Greenland did not indicate a large increase in Region 1/2 current. Inspection of polar cap data suggests some increase in the sunward perturbation near or just after the time of onset. It is uncertain to what extent this is due to the large increase in the current wedge westward current and its associated Region 1/2 currents. The increase in sunward perturbations in the polar cap during the CDAW 9C event discussed in Chapter 8 is small and suggests that these perturbations are primarily due to Region 1/2 currents with the already strong driven system. The same may be said for the Chapter 10 GEM event. To summarize, conclusions about the effects of onsets on the global Region 1/2 systems are difficult to make due to the sporadic satellite passes and the fact that their ground perturbations are relatively subtle. There are indications that they do not greatly respond to substorm onsets on a global scale, but react much as do the driven system electrojets. The only exception to this appeared to be in the case of the CDAW 9A substorm, in which the current in the substorm current wedge and its associated field-aligned current system dramatically exceeded that present before onset.

A final point about the global electrojets is that they do not generally appear to extend, on the dayside, to near the longitude of the subsolar point (that is to magnetic local noon). DPY activity does take place there and was seen (in the CDAW 9C event of Chapter 8). In no other event is there any strong indication that the electrojets extend very far onto the dayside, and for the GEM study, with the best global data, it is quite clear that they do not do so in any organized fashion. This result appears to contradict that of Hughes [1978] which arrived at a schematic view of electrojets extending from near-noon onto the nightside by modelling statistical meridian profiles from a chain at many MLTs sampled as the chain rotated. Hughes' study in effect pre-averaged data by using the latitude of a characteristic shift in the Y component as a reference point and obtaining individual meridian profile matches once such shifting had been done. The technique is also discussed in Section 4.3, where results for several local time sectors are shown. A repetition of Hughes' study using the AFM technique would eliminate the need to perform a latitude shift and results could be averaged as latitudes after inversion rather than as perturbations before inversion.

c. Other Aspects of Geomagnetic Phenomenology

In three of the five events studied, westward magnetic perturbations in the dusk sector were noted. In three of these, these perturbations were latitudinally confined and not able to be modelled even after inclusion of Region 1/2 currents in the model. Among these three events, that of August 8, 1997 had the weakest such perturbations, and they were partially explained by the confluence of the eastward and westward electrojets as a 'Harang discontinuity'. Even in this case, however, the perturbations were not completely matched by modelling, and in the other cases they were stronger. A suggested attack on the problem of dusk sector westward perturbations would be to use meridian chain data from two longitudes (CANOPUS and IMAGE are ideally placed), identifying onsets in the midnight sector at one chain while monitoring westward perturbations at the other. Such an approach should readily lead to the confirmation or refutation of the dusk sector westward perturbations which are reported here. This would be needed before proceeding to explain their cause.

d. Use of activity indices

It has been shown to be relatively easy to use automated fitting to model latitude profiles in regions where one electrojet is dominant. Cases have been shown where a very poor initial condition would lead to what appeared to be a reasonable final result as verified by manual inspection of the appropriate latitude profile and magnetograms. It was found through experience that a large value of the scale parameter for the Levenberg-Marquardt routine (typically 10) permitted such robust performance, whereas in a few cases when the wrong scale parameter was (usually inadvertently) used, poor convergence took place and sometimes even initially misleading results were produced. The routine presented here can thus be claimed to be a robust way of converting magnetic meridian chain data into information about the toroidal current across the meridian in question along with the basic geometric parameters of the electrojet. Even in the case of the simplest electrojet model, very good representation of the data was possible as quantified by accepted statistical measures of correlation.

In cases where the standard indices were available, these have generally been found to be lacking. The AE(12) index suffers from poor station coverage and may give misleading indications of time of onset due to longitudinal offsets. It was shown at the end of Chapter 6 that the index can miss events shown in the horizontal component even at stations relatively near an AE station in longitude, but at a different latitude. It has been pointed out [Rostoker, 1993] that, ironically, the tendency of the auroral oval to expand, so that activity is centred equatorward of the standard AE stations at times of high activity, can give rise to a decrease in the AE index when in reality the activity, by any realistic measure, increased. These undesirable effects can be remedied to a certain extent by the use of more stations in formulating an AE(n) index, and that has become a common practice as auroral data become available, quickly and on a regular basis, from more stations. The basic problem of perhaps missing onsets or misparameterizing them remains, although reduced in impact. We can not expect to fill the Earth limitlessly with magnetic observatories, and station gaps must continue to exist, especially in offshore regions

An attempt to remedy this can be made through use of more data than merely the horizontal intensity. Cases have been discussed in which the vertical component clearly showed perturbations which were not strong in the horizontal perturbations at any of the available stations. Certainly, the vertical component contains information which should be used, albeit carefully due to its sensitivity to local induction effects. Recently, the CANOPUS project has supplied meridian chain data which has been widely used, and informal use of indices based on this chain alone has become common. Rostoker has introduced the use of indices (CU and CL) defined as the envelope of total perturbation, and calculated as the extrema of $\mathrm{sign}(H)\sqrt{H^2 + Z^2}$, where H and Z are the respective horizontal and vertical perturbations, among stations along the CANOPUS Churchill meridian [Rostoker *et al.*, 1995]. A fitting routine also exists which can roughly determine the electrojet borders in near-real time, largely based on Z extrema. These measures go part way toward producing a useful indication of auroral oval magnetic activity. As mentioned in Chapter 11, the fact that these useful indices come from only one particular meridian can also give a misleading impression of global activity if they are used in isolation from other data.

What has been outlined in this thesis is a routine which is flexible enough to be used for modelling global current systems, but which shows its main strength in inverting magnetic data from meridian chains. Such inversions, with a simple model, give the total current and basic geometric parameters in a reliable and unbiased manner. Extension to cases where two electrojets in opposing directions cross a meridian is straightforward and has in fact been done although the results were not presented (although some of the three-dimensional models do feature this situation). With good latitudinal coverage the single number arising from such inversions would be total current through a meridian, which is a physical parameter as opposed to an indirect indication of magnetic perturbation as is the AE index. In cases where opposing electrojets occur, the two numbers for eastward and westward current would be available. Such a measure of current flow would be a more reliable indicator for correlative studies than would a magnetic index such as AE. Indeed, it would seem profitable to undertake such studies to establish relationships between total current and geometrical parameters, based on the relationship of electrojet width to onset which has been noted in the events examined here.

The present routine would also be well suited to repetition of surveys such as those of Rostoker and Phan [1986] or of Hughes [1978]. A very large data set could be treated in an automated fashion, allowing subdivision of the survey into finer gradations of activity than were used by those authors. Such a survey would give a better view of the distribution of auroral currents, with meaningful statistics and correlations to other studies being possible. A possible correlation could be envisaged, for example, with the eastward perturbation at geosynchronous orbit. The changes alluded to by Skone [1994] and briefly examined in Chapter 6, could be statistically firmed up by a correlative study of electrojet borders, current (which also gives an idea of onset times), and perturbations at GOES satellites, which are readily available. The concept of the Region 1/2 current system sweeping over equatorial satellites, and thus giving rise to changes in the eastward component, could readily be tested in such a manner, with obvious implications for mapping studies. A recent publication in a related vein is that of Lu *et al.* [1997], in which recent general magnetospheric models (T95 and T96, which are improvements on the widely used Tsyganenko fitted models) are used to provide good representations of the geosynchronous field as observed at GOES, but which appear to lack agreement in the details of the eastward field which may most plausibly be related to field-aligned current effects.

Studies utilizing three-dimensional modelling can also be envisaged although the problem of data sparsity would remain if direct use of observatory data is made. In principle, the approach of Hughes [1978] of assembling meridian profiles into a simulated global distribution of stations could solve this sparsity problem. It would be possible, using global AFM, to perform fits directly to rotated station data without such a shift, and after the fact to determine what levels of shifts in latitude occur in the real system. This approach may not be optimal since in this study the global models have proven difficult to use and to interpret. However, a study of Region 1/2 effects would in principle be possible using such an approach if polar cap stations were included.

e. Technical Considerations

In addition to investigation of magnetic field phenomenology associated with the auroras, this thesis has had as an intent the development of useful tools for such investigation. The Levenberg-Marquardt technique has been at the heart of the tools developed for magnetic inversion. It has been used in conjunction with the routines of Kisabeth [1972], slightly improved upon, with graphical display techniques, and with ancillary data, to assist in the interpretation of magnetic perturbations and to invert the information they contain into information about physical current systems. An assessment should now be made of the technical standing of this approach, that is, to what extent a useful tool has been developed and what could be done to improve it.

The assessment of utility depends largely on what one aims to use a tool for. If the intention here was to do global inversion of the magnetic perturbations into a global view of the current systems, then the routine can be consider to have worked well as a tool. In all cases for which this was tried, a final result was obtained which appeared to be generally reasonable in terms of the resulting current systems and which reproduced the input data to a satisfactory degree. In retrospect, however, much human intervention was required. The development of informative graphical displays allowed such intervention and in some cases led to useful insights while the modelling process was ongoing. For example, the 'northerly onset' found in Chapter 6 was recognized while trying to understand why the modelling routine placed a current system at an unusual location. The recognition that polar cap perturbations can be reproduced in many cases through the inclusion of Region 1/2 currents in the auroral oval was another way in which use of the routine was instructive. The fact that one obtains quantitative information about the current systems is also very useful: such information is likely to be superior to less physical parameters such as AE in making correlative studies.

The success in global modelling must be tempered by the fact that use of the routine in this maximal form is very difficult, both for the operator, and for the computer. Station sparsity is such that convergence may not be possible unless a reasonably good initial model is supplied. This can generally be done through analysis of intermediate output results, but the iterative process involved, with relatively slow convergence of the large global models, is onerous. Approaches involving a preliminary step to automatically set up crude initial models do not seem promising. In practice, initial models for global modelling were set up based on very good local inversions from meridian chains, yet they needed considerable later 'tweaking' to finally attain convergence. It also is the case that badly modelled stations can have their weight increased until the routine is forced to optimally use their information. Once more, one might naively think that this process, if even desirable, could be automated. However, judgment is essential in such reweighting, since stations may be affected by various sorts of noise. In judging the most profitable route to take with AFM, global modelling does not seem to recommend itself.

Fortunately, the approach taken in AFM is versatile. As opposed to methods based on global decomposition of the observed magnetic fields (usually through spherical harmonic analysis), AFM is local. Its use for global modelling depends on extension of many local cases to the global case, the coherence implied being supplied through use of large-scale current systems. But since AFM is local, it can be used regionally, or in principle even for a single station. As has been repeatedly shown in this study, regional use of AFM can be very successful in reproducing the input data with a convincing yet simple input model. A further advantage is that these inversions generally proceed automatically or nearly so. They are, in practice, far more robust than is global modelling.

There are several reasons for which regional analysis works better than does global analysis. One is that the density of data is inherently better: one selects regions where data exist, not regions without data! Clearly one has no choice in deciding to do global modelling: the whole globe is there and data are only where they have been measured and made available. The result is that regional modelling, which has less parameters involved yet is done in a data-rich area, has enough data to effectively drive those parameters through the many dimensional parameter space in which convergence is sought. With less parameters, gradients in the parameter space become steeper and easier to follow. Another aspect is that interpretation of the final parameters is more apparent when there are fewer of them: despite large efforts having gone into display of results from global modelling, it is not always easy to see subtle changes between time steps. The comparable information from regional modelling, the electrojet borders for example, are easily displayed on a simple graph.

Another technical aspect is the modelling technique at the heart of the routine. In this case, Kisabeth's method has been used. This method requires a large amount of expensive (of computing time) integration. It has the further disadvantage that it is a discretized integrator using Gaussian integration. While Gaussian integration is among the best for tabulated numbers at discrete points, the need to represent the currents at discrete points, when this is far from reality, can give problems with its use. When Kisabeth's expensive method is used with the Levenberg-Marquardt technique on a global scale, many partial derivatives must be taken, and the result is a cumbersome numerical technique. Use of a method involving linear currents rather than discrete ones, and an efficient method for obtaining partial derivatives, would improve the modelling efficiency. If use only on a local scale, with few parameters, is envisaged, then the technical development implied in the foregoing would not be called for and use of the existing routine would be practical.

f. Contextual Assessment of AFM

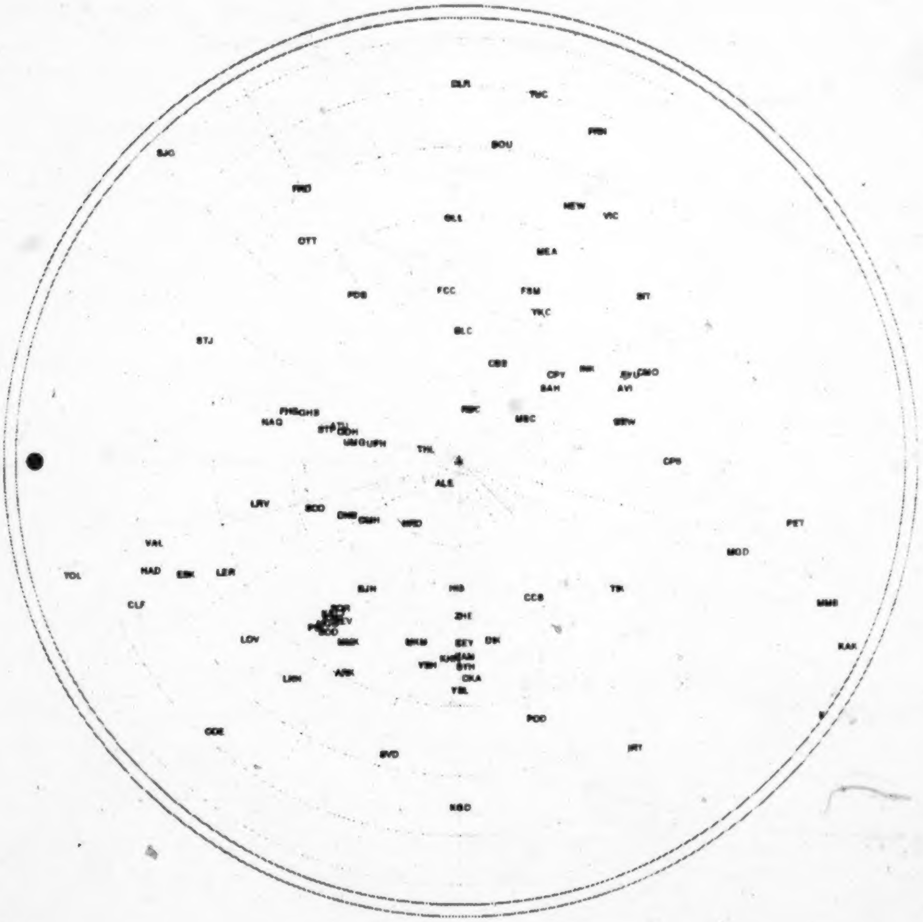
Having shown that the Automated Forward Modelling technique can invert surface magnetic observations into meaningful three-dimensional current systems, it is now useful to look at how well it does so in the context of existing techniques for such inversions. The most commonly cited model in the current literature is AMIE, which was discussed in detail in Section 4.c. AFM and AMIE results have been presented for comparison previously [Connors *et al.*, 1995] but here the two techniques are compared in a more quantitative and critical manner.

A complication in comparing results from AFM and AMIE is that the two routines use different coordinate systems. The centred dipole coordinates used by AFM were described in detail in Section 3.c. AMIE uses the Magnetic Apex Coordinates of Van Zandt *et al.* [1972] which are described in detail in the cited reference. For practical purposes it is noted here that the two systems may differ by several degrees, much as the centred dipole coordinates differ from corrected geomagnetic coordinates as illustrated in Figure 3.12. This difference is primarily evident in what follows as latitude differences and an attempt will be made to clarify apparent latitude differences arising from coordinate system effects from possibly real latitude differences in the output from the models from either method. In longitude, the difference is usually not sufficient to cause concern; AMIE results are presented in MLT while AFM results have been presented here in an Earth-fixed system but generally with the subsolar longitude near the top. Bearing in mind that in the MLT system the subsolar longitude is by definition at the top of the diagram, the longitudes may be readily compared.

Apart from the coordinate systems, a further obstacle to comparison reflects the difference in basic approach of the two modelling systems. AFM focuses on the integrated aspects of current systems and uses a simplistic systems of a small number (usually three) of electrojets, with associated field-aligned currents, to obtain a small number of parameters which optimally reproduce the input data and thus constitute a data inversion. AMIE, by contrast, uses a large number of parameters and represents physical quantities, including the parameters of electrojets and field-aligned currents, at many points on the Earth's surface (refer back to Section 4.c). Quantities such as total currents are best compared by integrating over grid points from an AMIE solution. Other quantities, such as geometric parameters, are best compared using graphical output from each routine, taking the coordinate system differences into account.

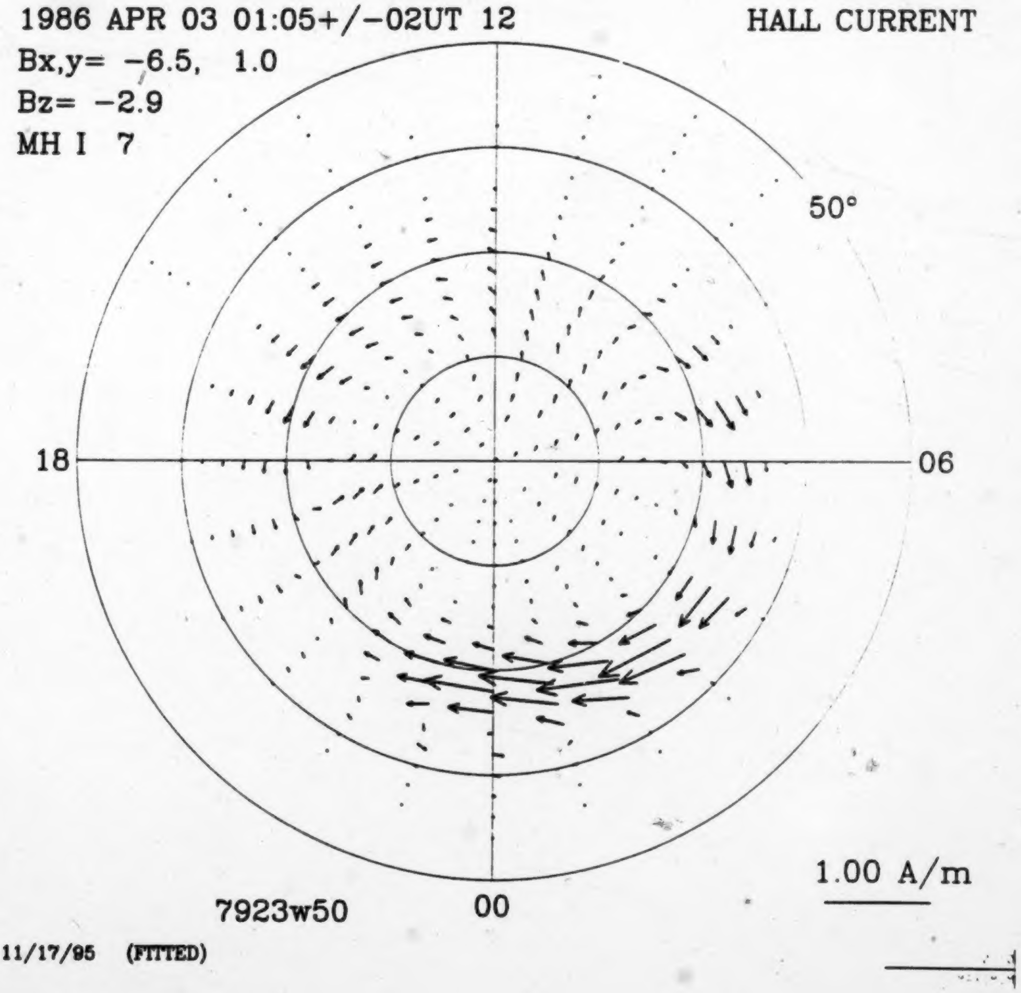
To assist with data comparisons, Figure 11.1 shows the positions of stations in Apex Magnetic Coordinates along with their names. This may be compared to Figure 6.25 showing most of the same stations in the centred dipole system.

Figure 11.1 Positions of stations available for AMIE modelling in Apex Magnetic Coordinates. Same stations apply for both April 1986 events described in the text but the orientation here is for April 1, 1986 at 1905 UT. For April 3, 1986, the black dot at left should be at bottom. In closely spaced regions, consult Figure 6.25 for names.



The first comparison which will be presented is for the CDAW 9B event of April 2-3, 1986, which was detailed in Chapter 6. This comparison is similar to that made by Connors *et al.* [1995] but more detailed. Figure 11.2 shows the Hall currents as calculated by AMIE shortly after the initial major onset on April 3, 1986, which took place at 0055 UT. The figure shows conditions at 0105 UT and may be compared to Figure 6.30 bearing in mind the points made above regarding coordinate systems. A qualitative comparison shows substantial agreement between the two figures. There is substantial current flow only in the midmorning to immediate pre-midnight local times in both figures. The AFM result assigns this to the morning electrojet, which extends past midnight, and the current wedge which is from about 0400 to about 2300 MLT. AMIE does not separate out current systems in this manner but the zone carrying enhanced current, which could most reasonably be assigned to the wedge, spans a similar MLT range in the AMIE result.

Figure 11.2 AMIE output showing Hall current on April 3, 1986, at 0105 UT. Scale for 1.00 A/m is shown at lower right. Courtesy of Barbara Emery.



The general appearance of the zones of westward current flow is similar, with the morning electrojet from about 0800 to 0400 MLT flowing at roughly constant latitude, while the current wedge rises in latitude as it approaches and passes midnight. In the AFM result, the wedge is accompanied by an extension of the morning sector electrojet, and arguments for the presence of this doubled electrojet configuration have been given in Chapter 6. AMIE does not indicate such a doubling, but has a single band of westward current which is enhanced towards its centre. In the evening sector, both methods suggest a weak eastward electrojet from about 1500 to about 1900 MLT, with a decrease in latitude as midnight is approached. Both methods also indicate a zone relatively devoid of meridional current flow in the mid-evening sector: as discussed in Chapter 6, there is a gap between the evening sector and morning sector/wedge electrojets. The AMIE result supports the contention that there was little meridional current flow on the dayside, which was stated in Chapter 6. Finally, the AMIE result suggests little Hall current to be flowing in the polar cap. The modelling results for the 0055 onset in Chapter 6 did not include Region 1/2 currents and their ionospheric closure, and it was later shown that their presence is needed to adequately model polar cap perturbations. In Figure 6.30, it can be seen that these perturbations were not well modelled. The lack of Hall current in the polar cap supports this conclusion: presumably AMIE modelled the polar cap perturbations with substantial Pedersen currents which it indicated to be essentially colocated with the wedge and morning sector electrojets. Since this point is not essential to comparison of AMIE and AFM results, these currents are not illustrated here.

To some degree, quantitative comparison of the results of the two methods is possible based on Figure 11.2. By measuring the lengths of arrows indicating Hall current flow, one can integrate the Hall current across critical meridians for comparison with AFM results. As a point of technique, the original scale bar on AMIE output was 18 mm, and total current flow, in MA, across a meridian could be computed as the total length of arrows showing appreciable perturbations multiplied by 0.0123. From AFM output listings, the total current in any given system could be directly read off, and used at the meridians which it crossed. Since both the wedge and morning sector electrojets crossed the midnight meridian, the current from these two systems was summed there. The quantitative results of this comparison are shown in Table 11.1

Table 11.1 Quantitative comparison of current across meridians in AMIE and AFM on April 3, 1986, at 0105 UT. Geometrical parameters for AMIE are in Apex Magnetic Coordinates, for AFM in the centred dipole system.

Meridian (MLT)	System	AMIE			AFM		
		MA	Width	Centre	MA	Width	Centre
00	Wedge	0.43	5°	69°	0.40	3°	71.5°
06	Morning	0.19	4°	68°	0.11	3°	64.5°
18	Evening	0.091	4°	68°	0.098	3°	65.5°

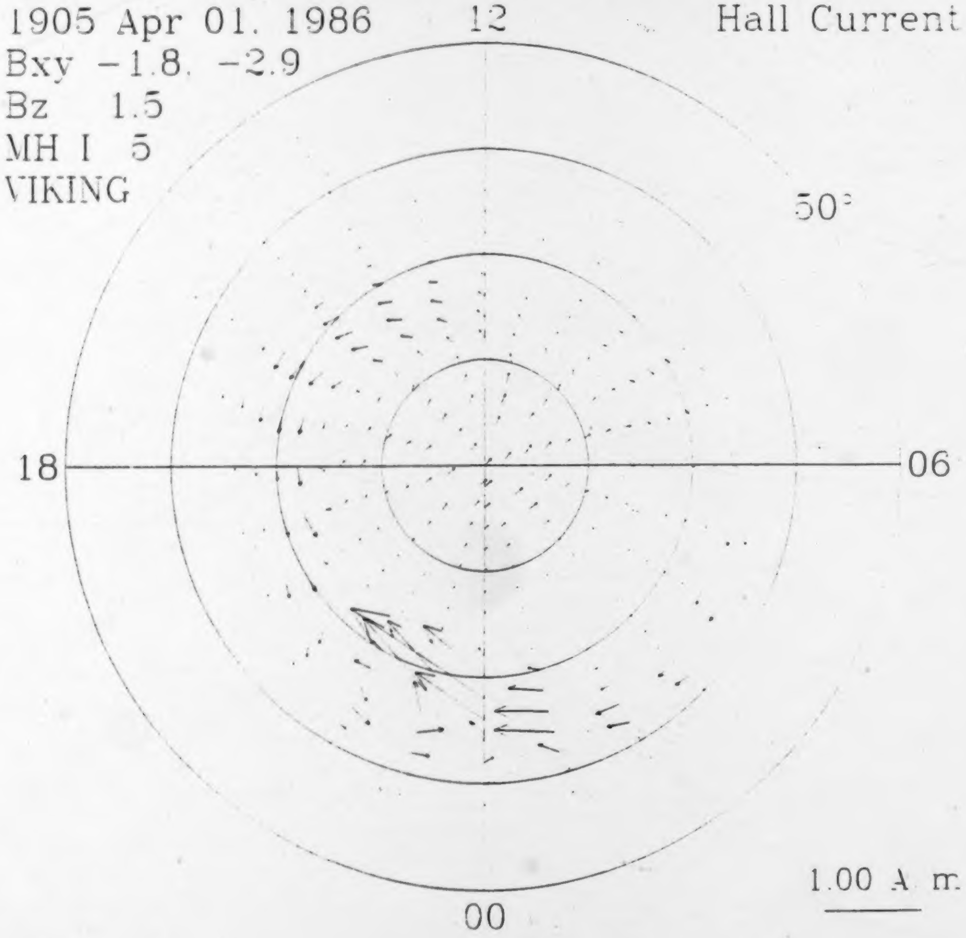
The current crossing midnight (due to the wedge or wedge/morning current systems) is seen to agree well between the two solutions, while AFM indicates less current at 0600

MLT than does AMIE. The weak evening sector electrojet is indicated as similar by both solutions. Widths of the electrojets are slightly larger in AMIE than in AFM. It appears initially as if there is a large divergence in latitudes indicated by the two solutions. This is in fact an artifact of the coordinate systems used. This may be seen by comparing the latitudes of stations near the various meridians, in Apex from a printout supplied with AMIE results [Barbara Emery, private communication, 1998], and in centred dipole (CD) from Table 6.8. The midnight meridian is near NAQ and the Apex latitude there is 67.1°, while the CD latitude is 70.7°. The current flow centre is thus about 1° north of this station in both cases. Similar considerations hold for DIK in the morning sector (near the centre of the electrojet in either system), and in the evening sector where the AMIE electrojet is centred about 1.5° south of FCC and that found by AFM at GIL, i.e. the same latitude (GIL was not used in AMIE modelling). In a quantitative fashion, it is seen that there is good agreement between AMIE and AFM with only one exception. That exception is the strength of the morning sector electrojet. Inspection of Figure 6.30 will show that perturbations at stations near the northern border of this electrojet are not matched well in the AFM and that this is due to the electrojet being located south of them. It is likely that the electrojet should be wider, thus covering these stations. With its larger number of degrees of freedom, AMIE has been able to effectively produce perturbations at these stations whereas AFM has not been able to do so. However, this is only one aspect, and overall, the degree of agreement between the two independent solutions is considerable.

A yet more useful comparison may be made for the CDAW 9A event of April 1, 1986, which has been discussed by several authors as described in the introduction to Chapter 7. Quantitative details of AMIE runs corresponding to those published by Lu *et al.* [1997] have been provided by Barbara Emery [personal communication, 1998], allowing comparison of the AFM and AMIE routines to be better quantified in terms of their ability to represent input data. For this event it is also possible to make qualitative comparisons to the KRM technique as presented by Ahn *et al.* [1995]. Figure 11.3 shows the Hall current from an AMIE run done based on conductivity derived from a Viking image. This is considered to be the most appropriate conductivity model possible as the onset region was well imaged for the event (dayside conductivities are generic). This figure is most directly comparable to Figure 7.20 and a qualitative comparison will now be given.

The main feature of both figures is the strong substorm current wedge which crosses the midnight meridian. In both solutions, the wedge currents flow in the ionosphere for roughly one hour of local time on either side of midnight. The primarily meridional currents are angled poleward into the evening sector in both solutions. This is clear in the AFM output and is indicated both by the directions of arrows indicating local current flow in AMIE and by the trend of the longest arrows at each meridian. The north-south extent of these wedge systems is about 7° in each case as they cross midnight. The morning sector electrojet in each case is generally at lower latitude than is the current wedge, and arises at roughly the dawn meridian. The AFM solution has this system slightly tipped southward as the current flows toward midnight, a result largely dictated by perturbations at stations BRW and CPS in the midmorning sector. The AFM solution has a morning sector electrojet which approaches 10° in north-south extent as midnight is approached, whereas the AMIE solution has this electrojet flowing at nearly constant latitude with a width of about 4°. In both cases, the morning sector westward electrojet is considerably weaker than the current wedge. Both solutions have a weak evening sector electrojet at relatively high latitude (roughly 70°), with a slight tendency to decline in latitude as midnight is approached. In each case this system has north-south extent of about 5°, starts at about noon, and flows toward the general region of the wedge. In the AMIE solution there is an eastward electrojet south of the wedge current but it appears to be disconnected from the primary evening sector electrojet. In the AFM solution the eastward electrojet was constrained to intersect the wedge in a common region of upward field-aligned current. In both solutions there is little current flow in the dayside morning sector, nor across the polar cap.

Figure 11.3 AMIE output showing Hall current on April 1, 1986, at 1905 UT. Scale for 1.00 A/m is shown at lower right. Courtesy of Barbara Emery.



01/23/98

A more quantitative comparison will now be attempted. Much as was done above, the lengths of Hall current density vectors as plotted in Figure 10.3 were measured. The original length scale was 17 mm long, so that a conversion factor of 0.0131 was used to convert the total length of such vectors crossing a meridian into the current in MA flowing across it. The original program listing used in preparing Figure 7.20 was consulted and the current associated with each system determined. The results are shown in Table 11.2.

Table 11.2 Quantitative comparison of current across meridians in AMIE and AFM on April 1, 1986, at 1905 UT. Geometrical parameters for AMIE are in Apex Magnetic Coordinates, for AFM in the centred dipole system.

Meridian (MLT)	System	AMIE			AFM		
		MA	Width	Centre	MA	Width	Centre
00	Wedge	0.82	7°	68°	0.51	7°	66°
03	Morning	0.12	4°	63°	0.18	10°	65°
18	Evening	0.10	5°	71°	0.30	5°	73°

These results do not appear to be as concordant as were those for the April 3 1986 event described above. Before proceeding to attempt to explain the differences it is useful to consult Figure 11.4, which graphically shows what AMIE calculated at each station based on its solution, along with the original data. This figure presents the perturbations in a form similar to that used to show observed and modelled perturbations in AFM, and may be directly compared to Figure 7.20. It is immediately clear that AMIE has done an excellent job of matching the perturbations used as input. Much the same could be said of AFM based upon inspection of Figure 7.20. A quantitative discussion of the match to input data is reserved for later. For the moment, the figures, along with the station key of Figure 11.1, will aid in understanding the differences which are apparent in Table 11.2.

The first apparent difference is in the total current of the substorm current wedge. In the midnight meridian, AMIE appears to indicate that the wedge is stronger, at 0.82 MA, than does AFM, with 0.51 MA. This is at least in part due to the small longitudinal size of the current wedge. Measurement of the AMIE vectors shows that the 01 MLT meridian had only 0.40 MA crossing it, while the 23 MLT meridian saw 0.44 MA. The average current in the vicinity of the midnight meridian as shown by AMIE is thus less than is that in the midnight meridian itself. The disagreement is considerably less than it initially appears. The widths from the two systems are in good agreement, and it may be seen by inspection of the figures that the latitudes in the respective current systems indicate that the wedge stretched from about the middle of the European Russian set of stations (i.e. near TAM) to BJN north of Scandinavia. Thus the two solutions do agree about basic characteristics of the wedge.

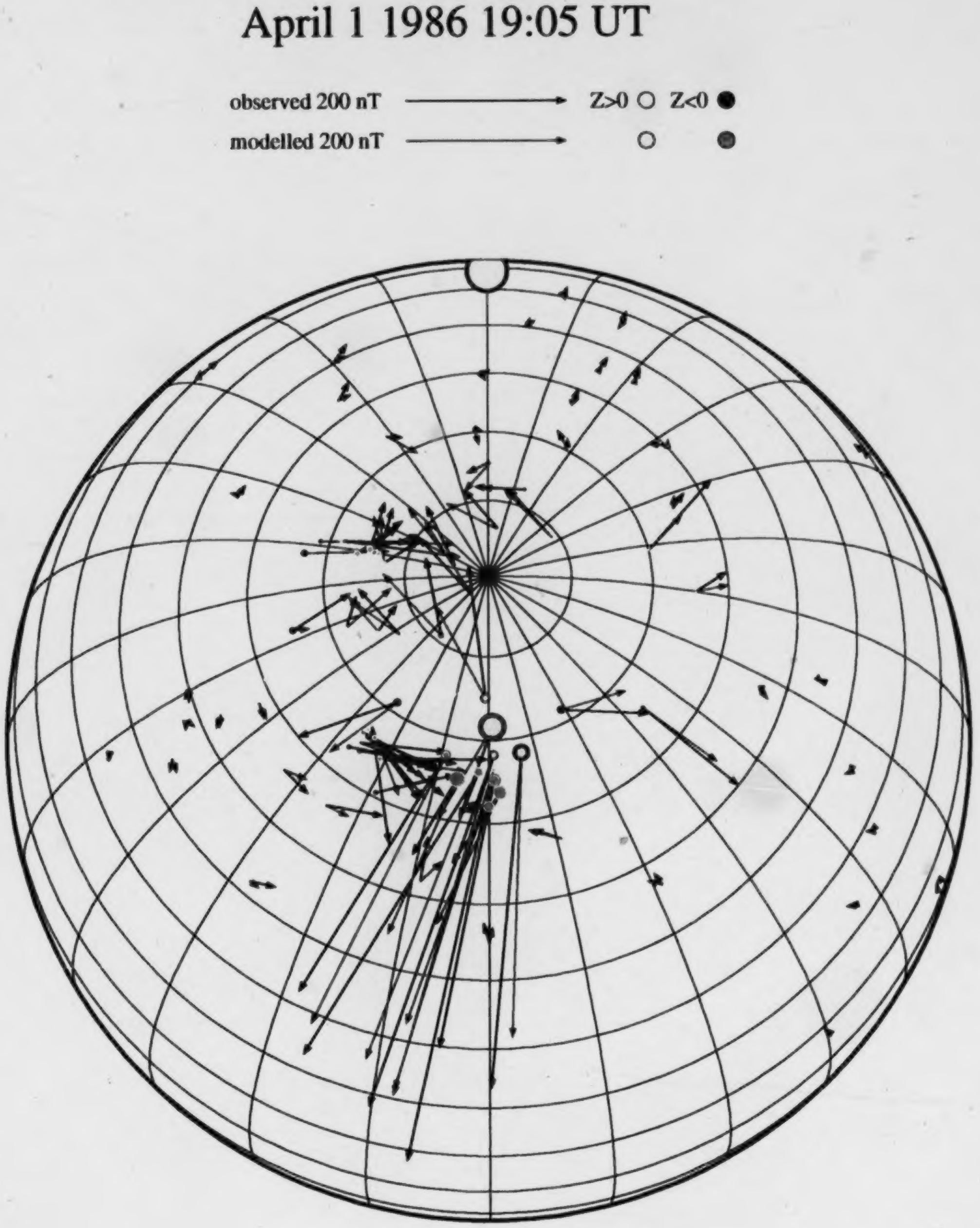


Figure 11.4 AMIE model and computed perturbations in Apex coordinates. Result is based on conductivities derived from Viking imaging.

The situation appears to be reversed for the morning sector electrojet: AMIE indicates it to carry less current than does AFM, with 0.12 and 0.18 MA, respectively. The width found by AMIE is less than that found by AFM, and inspection of the relative locations of CCS and TIK and the electrojet shows that AMIE obtains a considerably more southerly latitudinal placement. Comparing Figures 11.4 and 7.20, reference to CCS and more particularly to BRW further into the morning sector suggests that AFM has found a more realistic configuration of the electrojet, since those stations are better represented in it, yet not at the expense of a match elsewhere.

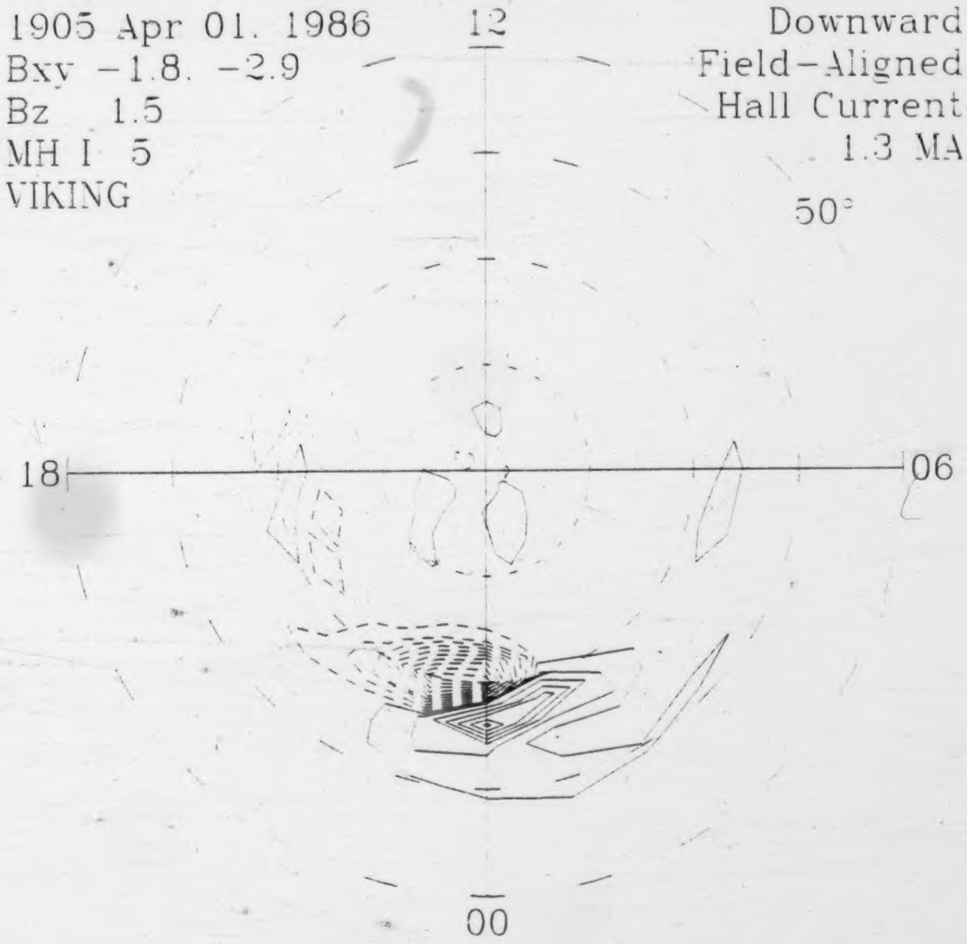
In the 1800 MLT meridian, the underspecification of current by AMIE initially appears to be repeated, since it shows only 0.10 MA, while AFM shows 0.30 MA, crossing that meridian. The width and positioning of the electrojet may be inspected and found to be very comparable, so only the current needs discussion. Once more, there appears to be longitudinal variation in the AMIE solution, yet this time in a region which is not as active on a small scale as that near the wedge. In Figure 11.4 it should be noted that the station LRV, which is near the 1800 MLT meridian, is very poorly represented, with far less modelled perturbation than observed. The electrojet as depicted by easterly directed Hall current arrows in Figure 11.3 appears to be diminishing in the 1800-1900 MLT range. In the 1600 MLT meridian, there is total Hall current flow, as calculated by AMIE, of 0.20 MA. Perturbations in this MLT range seem better modelled, for example at FHB there was 120 nT of X computed versus 108 nT observed. At NAQ, however, the corresponding numbers were 76 and 105 nT. The case can certainly be made that AMIE underrepresented the current through the 1800 meridian, and perhaps got a better value in the 1600 MLT meridian. AFM appears to have succeeded better on average in this MLT range than did AMIE.

The above considerations of ionospheric electric current flow have led to the conclusion that there is substantial agreement between AMIE results and those of AFM, particularly when variations in MLT within AMIE are accounted for. These variations are smoothed out in AFM through its use of longitudinally continuous current systems, whereas those of AMIE are permitted more variation with this coordinate. Given the station distribution available, it is not possible to say which approach better represents reality.

Since this AFM run, unlike the Apr 3 case discussed above, included field-aligned currents, it is of interest to attempt to compare these between the two modelling techniques. This task is more difficult than is the comparison of ionospheric currents. In AFM, the locations of field-aligned currents (FAC) associated with basically meridional currents are shown on the output diagrams and the total current in FAC is the same as that in the ionospheric system. The Region 1/2 currents are restricted to have their ionospheric closure in a north-south direction and are colocated with the meridional currents which are more constrained by ground perturbations near the electrojets. It has been shown, in several cases above, that the Region 1/2 currents primarily explain polar cap perturbations when solved for with AFM and the simple three component current system which has been used. The treatment of field-aligned currents is quite different in AMIE or KRM. In these routines, the field-aligned currents are represented on a grid much as every other variable.

They are generally computed from the divergence of the ionospheric currents [Lu *et al.*, 1997]. The Hall and Pedersen currents may be separated in AMIE or KRM solutions, so that FAC associated with each may be specified individually. Figure 11.5 shows the FAC associated with the Hall currents discussed above. The figure of 1.3 MA shown at the upper right is the hemispherically integrated downward Hall-associated FAC. Since there are other zones of FAC besides those near the wedge, this figure is not inconsistent with the 0.82 MA indicated to be present as Hall current in the wedge in Table 11.2. The direction from the downward FAC region to the upward in the vicinity of the wedge is toward the northwest, which is consistent with the conclusions reached above about the Hall current flow in both AMIE and AFM. Since the Hall current in AFM was found to be only 0.51 MA in the wedge, the AMIE result introduces more FAC than does AFM. This is a natural consequence of AMIE's more concentrated wedge.

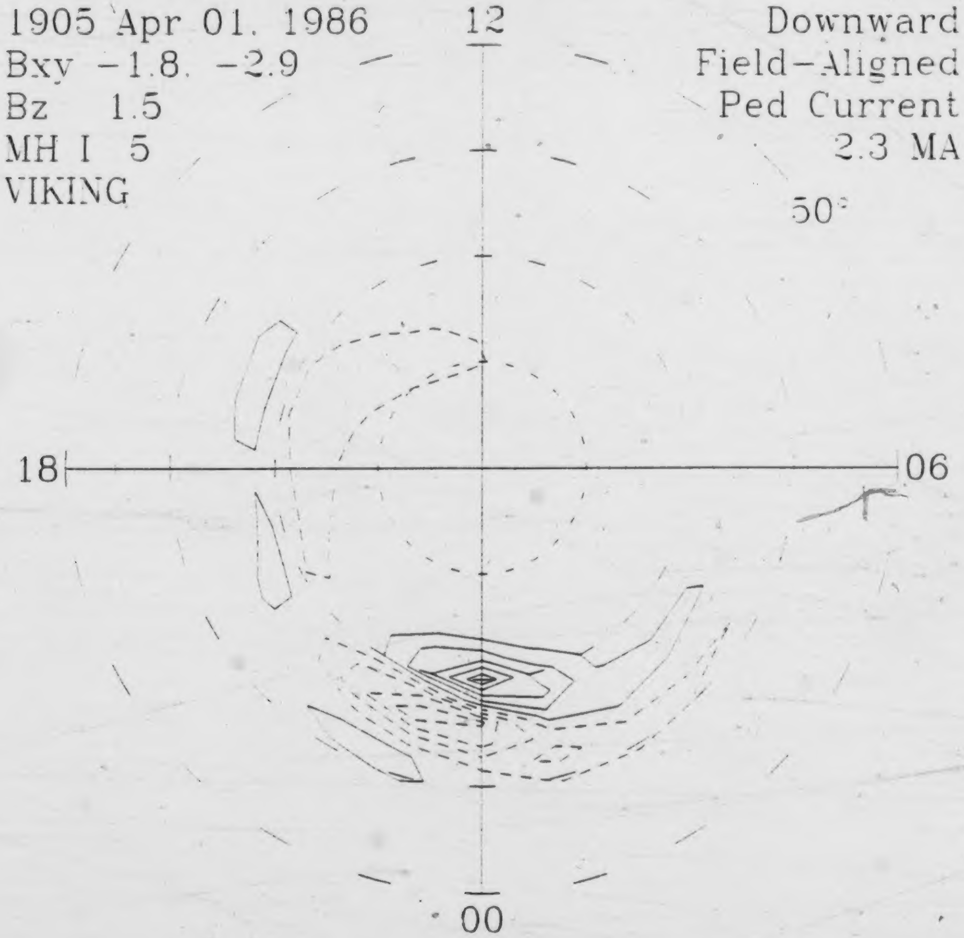
Figure 11.5 FAC associated with Hall current on April 1, 1986, at 1905 UT. Contours are solid for downward FAC, dashed for upward FAC, and contour interval is 0.15 $\mu A/m^2$. AMIE solution courtesy of Barbara Emery.



01/23/98

The Pedersen current found by the AMIE routine is shown in Figure 11.6. The lines joining downward and upward FAC regions would be essentially perpendicular to the Hall flow inferred above, a not unreasonable result. There appears to be about 2 MA of FAC concentrated mainly in the region near the wedge. There is a much more diffuse region of paired FAC in the evening sector which implies local northward Pedersen flow as would be expected there. This region may be estimated to roughly account for .3 MA, consistent with about 2 being in the wedge region. In AFM, the Pedersen currents in evening and wedge (taking account of the colocation rule in the code) are approximately colocated with the evening and wedge FAC indicated by AMIE. However, the wedge FAC from AFM is only about 0.72 MA, while that in the evening sector totals 0.79 MA. A further difference is that AMIE indicates little FAC in the morning sector. AFM, on the other hand, finds that as its most important FAC, at 1.5 MA total current.

Figure 11.6 FAC associated with Pedersen current on April 1, 1986, at 1905 UT. Contours are solid for downward FAC, dashed for upward FAC, and contour interval is 0.3 μA/m^2. AMIE solution courtesy of Barbara Emery.

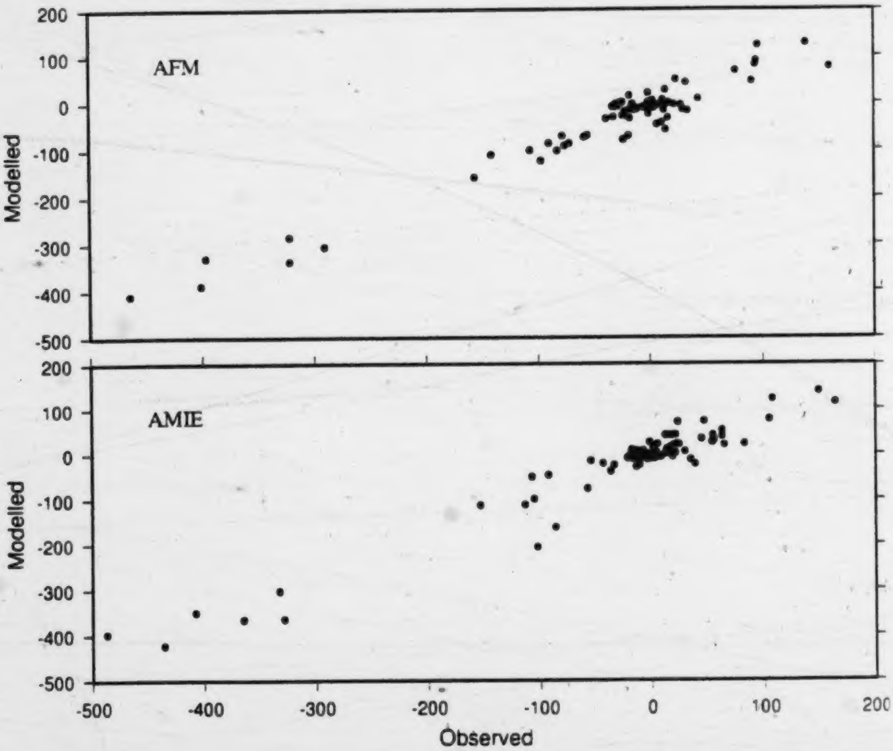


01/23/98

The total Pedersen-associated FAC in AFM thus seems to exceed that in AMIE and to have a different distribution. Comparison of Figures 7.20 and 11.4 shows that both AFM and AMIE reproduce well the polar cap perturbations. Given the low sensitivity of ground observations to paired FAC such as those associated with Region 1/2, the discrepancy may be difficult to resolve. One thing which might be said in favor of AFM's result is that it was shown in Section 7.c, through direct measurements by DMSP, that the morning sector FAC increased after onset. The AFM solution, with large FAC in the morning sector, would be consistent with this observation, whereas the AMIE result does not appear to be. To put things in perspective, though, the satellite measurements were taken well after (about one hour after) the time in question.

Having found Hall current effects in both AMIE and AFM to be in good agreement while results for Pedersen currents were not, it is now useful to take another tack in assessing the two routines: which represents ground measurements better? Figures 7.20 and 11.4 visually show that both techniques do a very good job of making the match of model results to data. Following the approach outlined in Section 9.c (see in particular Figure 9.13), we will plot data and model result for the perturbations and examine the correlation between them. This may be expected to give a quantitative assessment of the 'goodness of fit' of model to data for each technique. Figure 11.7 illustrates the basic idea using results of model fits to the X component using AMIE and AFM for April 1, 1986 at 1905 UT.

Figure 11.7 Observed and modelled X component values from AFM and AMIE for April 1, 1986 at 1905 UT.



In each panel of Figure 11.7, the observed X perturbation (in Apex coordinates for AMIE and in CD coordinates for AFM) is shown as the abscissa value, paired with the model value as the ordinate. In this manner one point is obtained for each station used in the model (79 for AFM and 82 for AMIE: missing stations differ for each program but most stations are common as may be seen by inspection of Figure 7.20 and Figure 11.4). In the case of perfect representation of the data by the model, a straight line with unit slope would be the regression line for each scatterplot, and the correlation coefficient r would be 1.00. It will be attempted to use the nearness of approach to this ideal situation to assess the relative merits of AFM and AMIE in matching observations through the model. Under two headings of AMIE and AFM, the r coefficient and the slope of the fit are presented in Table 11.3 for each magnetic component. It is noted that the r coefficients are all above 0.9, indicative of a strong positive correlation between the input data and the results calculated from the model. This correlation is stronger for the horizontal components and slightly weaker for the Z component, quite in agreement with comments throughout the thesis about the sensitivity of that component to small errors in electrojet placement. Both AFM and AMIE appear to underpredict variations in perturbations modelled as compared to observation, as indicated by the slopes of the modelled versus observed regressions being less than one. Only here does AFM have a coefficient which indicates poorer performance than does AMIE: the X variations are underpredicted by AFM about 1% worse than by AMIE. That AFM predicts Z better than does AMIE tends to reinforce points made previously about the good latitudinal response of AFM as it is this response which permits better modelling of the vertical component.

Table 11.3 Assessment of AMIE and AFM with regression parameters

Coefficient	AMIE			AFM		
	X	Y	Z	X	Y	Z
r	0.9687	0.9549	0.9157	0.9699	0.9608	0.9247
slope	0.9016	0.8444	0.8230	0.8914	0.8990	0.9151

Apart from the slightly better response of AFM to the Z component, the striking aspect of this table is not differences in ability to model data accurately, but rather that the two techniques have a good and similar ability to do so. Differentiating between them would have to be done on some other criterion in practice. This criterion could be ease of use, in which case AMIE would likely come out ahead as it is a well-developed procedure. It could be the ability to use only magnetic data to perform the inversion, in which case AFM would be preferred since it is a Biot-Savart solver making minimal assumptions about other quantities. It could be ease of interpretation of results, but this to a certain degree depends upon one's way of viewing results. Certainly AFM proposes a simple model view in terms of three basic current systems and attains a surprising degree of success in parametrizing geomagnetic data in those terms. However, AMIE may present the global distribution of current flow more accurately. Perhaps it is best to view the agreement between AMIE and AFM as mutual validations, leaving each program to be used in domains where it is most appropriate.

g. Suggestions for Further Work

Having validated the AFM technique by finding its results to be generally reasonable and in certain cases quite comparable to those of AMIE, it is appropriate at this point to consider what use it may have in the future.

As has been made clear above, the most promising domain for use of the AFM technique is regional or meridian chain work. Several ideas have already been mentioned involving mapping global systems using rotating meridian chains, studying Region 1/2 current movement in conjunction with geosynchronous satellites, and use of two chains to study events at or near the time of substorm onset.

A project to which AFM could be well suited would be electrojet studies on subcontinental scale, following on from work with the Scandinavian Magnetometer Array done in the late 1970s and early 1980s [Küppers *et al.*, 1979]. Sample use of AFM in this study which best illustrates the approach is the detection of Region 1/2 current over Greenland during the CDAW 9A event described in Chapter 7. A perturbation in the eastward component was detected which could be well explained in terms of the Region 1/2 system linked through the ionosphere. This approach could be carried further if, on a regional scale, the local tilt of the electrojet could be determined. Within a two-dimensional array of stations this may be readily done with AFM. The contributions to the eastward component from the electrojet and the coupled Region 1/2 currents could be simultaneously modelled, that is to say, the currents in the two systems simultaneously determined. After calibration through use of satellite data, AFM could routinely produce physical parameters of the electrojet which would be of use in applications in Space Weather and in correlative studies to better our understanding of the Solar-Terrestrial Interaction.

h. Synopsis

In this thesis, a new tool, Automated Forward Modelling (AFM) has been developed and validated. It has proven capable of characterizing global or regional electric current systems associated with auroral activity. It has been shown to produce results which are reasonable in light of auroral phenomenology presented for the five event sequences on which it was tested. It was shown above that for two of these events the results may be compared to results from AMIE and give a similar view of the auroral currents during active times, with similar high capability of matching input data. AMIE and other existing techniques offer a more general view of global auroral electrodynamics in that they deal with more physical parameters. This can be a weakness in that not all inputs may be readily available. AFM, on the other hand, works using only magnetic perturbations as input and may be used regionally to good effect. Thus the new routine complements existing ones. A particular strength is that AFM offers a limited number of readily visualized parameters which may be applied for analysis, in the true sense of the word, of the magnetosphere. Directions for such analysis have been laid out and AFM stands ready for application.

References

Abramowitz, M., and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, 1965.

Acton, F. S., *Numerical Methods that Work*, Mathematical Society of America, Washington, 1990.

Ahn, B.-H., Y. Kamide, H. W. Kroehl, M. Candidi, and J. S. Murphree, Substorm Changes of the Electrodynamic Quantities in the Polar Ionosphere: CDAW 9, *J. Geophys. Res. 100*, 23845-23856, 1995.

Akasofu, S.-I., The Development of Auroral Substorm, *Planet. Space Sci.*, *12*, 273-282, 1964.

Akasofu, S.-I., *Polar and Magnetospheric Substorms*, D. Reidel, Dordrecht, 1968.

Akasofu, S.-I., The Magnetospheric Currents: An Introduction, in *Magnetospheric Currents*, edited by T. A. Potemra, American Geophysical Union, Washington, pp. 29-48, 1983.

Akasofu, S.-I., Development of Magnetospheric Physics, in *Magnetospheric Substorms*, edited by J. R. Kan, T. A. Potemra, S. Kokobun, and T. Iijima, American Geophysical Union, Washington, pp. 3-9, 1991.

Akasofu, S.-I., Assessing the Magnetic Reconnection Paradigm, *EOS Trans. American Geophysical Union 75*, pp. 249,252, 1994.

Anger, C. D., J. S. Murphree, A. Vallance Jones, R. A. King, A. L. Broadfoot, L. L. Cogger, F. Creutzberg, R. L. Gattinger, G. Gustafsson, F. R. Harris, J. W. Haslett, E. J. Llewellyn, J. C. McConnell, D. J. McEwen, E. H. Richardson, G. Rostoker, B. R. Sandel, G. G. Shepherd, D. Venkatesan, D. D. Wallis, and G. Witt, Scientific Results from the Viking Ultraviolet Imager: An Introduction, *Geophys. Res. Lett. 14*, 383-386, 1987a.

Anger, C. D., S. K. Babey, A. Lyle Broadfoot, R. G. Brown, L. L. Cogger, R. Gattinger, J. W. Haslett, R. A. King, D. J. McEwen, J. S. Murphree, E. H. Richardson, B. R. Sandel, K. Smith, and A. Vallance Jones, An Ultraviolet Auroral Imager for the Viking Spacecraft, *Geophys. Res. Lett. 14*, 387-390, 1987b.

Anton, H., Elementary Linear Algebra, Fifth Edition, John Wiley, New York, 1987.

Armstrong, J. C., and A. J. Zmuda, Field-Aligned Current at 1100 km in the Auroral Region Measured by Satellite, *J. Geophys. Res. 75*, 7122-7127, 1970.

Armstrong, J. C., S.-I. Akasofu, and G. Rostoker, A Comparison of Satellite Observations of Birkeland Currents with Ground Observations of Visible Aurora and Ionospheric Currents, *J. Geophys. Res. 80*, 575-586, 1975.

Ashour, A. A., The evaluation of the field of the currents induced in the Earth by an external field whose distribution is known numerically, *Radio Sci. 6*, 171-173, 1971.

Baker, D. N., T. I. Pulkkinen, R. L. McPherron, J. D. Craven, L. A. Frank, R. D. Elphinstone, J. S. Murphree, J. F. Fennell, R. E. Lopez, and T. Nagai, CDAW 9 Analysis of Magnetospheric Events on May 3, 1986: Event C, *J. Geophys. Res. 98*, 3815-3834, 1993.

Baker, D. N., P. R. Higbie, R. D. Belian, E. W. Hones, Jr., and R. W. Klebesadel, The Los Alamos Synchronous Orbit Data Set, in *The IMS Source Book*, edited by C. T. Russell and D. J. Southwood, American Geophysical Union, Washington, pp. 82-90, 1982.

Bargatze, L. F., D. N. Baker, R. L. McPherron, and E. W. Hones, Jr., Magnetospheric Impulse Response for Many Levels of Geomagnetic Activity, *J. Geophys. Res. 90*, 6387-6394, 1993.

Bates, D. M., and D. G. Watts, *Nonlinear Regression Analysis and its Applications*, Wiley, New York, 1988.

Baumjohann, W., Merits and Limitations of the Use of Geomagnetic Indices in Solar Wind-Magnetosphere Coupling Studies, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, Terra, Tokyo, 3-15, 1986.

Baumjohann, W., J. Untiedt, and R. A. Greenwald, Joint Two-Dimensional Observations of Ground Magnetic and Ionospheric Electric Fields Associated with Auroral Zone Currents 1. Three-Dimensional Current Flows Associated with a Substorm-Intensified Eastward Electrojet, *J. Geophys. Res. 85*, 1963-1978, 1980.

Baumjohann, W., R. J. Pellinen, H. J. Opgenoorth, and E. Nielsen, Joint Two-Dimensional Observations of Ground Magnetic and Ionospheric Electric Fields Associated with Auroral Zone Currents: Current Systems Associated with Local Auroral Break-ups, *Planet. Space Sci. 29*, 431-447, 1981.

Besprozvannaya, A. S., B. D. Bolotinskaya, A. L. Kotikov, E. I. Nemtsova, O. M. Pirog, O. A. Troshichev, A. V. Frank-Kamenetsky, and T. I. Shchuka, Structure of Auroral Zone Phenomena from the Data of Meridional Chains of Stations: the Lower Ionosphere during the Expansion Phase of Magnetospheric Substorms, *J. Atmos. Terr. Physics 53*, 275-279, 1991.

Boteler, D. H., and G. Jansen van Beek, in *Solar-Terrestrial Predictions - IV*, Volume 3, edited by J. Hruska, M. A. Shea, D. F. Smart, and G. Heckman, NOAA, Boulder, pp. 57-70, 1993.

Brathwaite, K. S., and G. Rostoker, DP2 Current System in the Ionosphere and Magnetosphere, *Planet. Space Sci. 29*, 485-494, 1981.

Buchert, S., W. Baumjohann, G. Haerendel, C. la Hoz, and H. Lühr, Magnetometer and Incoherent Scatter Observations of an Intense Ps 6 Pulsation Event, *J. Atmos. Terrestr. Phys. 50*, 357-367, 1988.

Burrows, W. E., Exploring Space, Random House, New York, 1990.

Carlson, H. C., Jr., and A. Egeland, The Aurora and the Auroral Ionosphere, in *Introduction to Space Physics*, edited by M. G. Kivelson and C. T. Russell, Cambridge, New York, pp. 459-502, 1995.

Chamberlain, J. W., *Motion of Charged Particles in the Earth's Magnetic Field*, Gordon and Breach, New York, 38 pp., 1964.

Chapman, S., Perspective, in *Physics of Geomagnetic Phenomena*, edited by S. Matsushita and W. H. Campbell, Academic, New York, pp. 3-28, 1967.

Chapman, S., and J. Bartels, *Geomagnetism (2 volumes)*, Oxford University Press, London, 1940.

Chen, A. J., and G. Rostoker, Auroral-Polar Currents During Periods of Moderate Magnetospheric Activity, *Planet. Space Sci. 22*, 1101-1115, 1974.

Clauer, C. R., and Kamide, Y., DP 1 and DP 2 Current Systems for the March 22, 1979 Substorms, *J. Geophys. Res. 90*, 1343-1354, 1985.

Clauer, C. R., and R. L. McPherron, Mapping the Local Time-Universal Time Development of Magnetospheric Substorms Using Mid-Latitude Magnetic Observations, *J. Geophys.Res. 79*, 2811-2820, 1974.

Connors, M., Wedge!, poster presented at the CNSR meeting Substorm!, University of Alberta, October 1993.

M. Connors, B. Emery, and G. Rostoker, Automated Forward Modelling and AMIE Applied to the CDAW 9B Period, *EOS, Transactions of the American Geophysical Union* **76**, F506, 1995.

Connors, M., and G. Rostoker, Source Mechanisms for Morning Auroral Features, *Geophys. Res. Lett. 20*, 1535-1538, 1993.

Connors, M., Rostoker, G., and Korotova, G., Modeling of the Evolution of Substorm Expansive Phase Currents, *EOS, Transactions of the American Geophysical Union* **72**, No. 44 (Supplement), 396, 1991.

Cramoysan, M., and D. Orr, The Use of a Model Current Wedge in the Determination of the Position of Substorm Current Systems, *preprint*, University of York, York, U. K., 1993. later published as Cramoysan, M., R. Bunting and D. Orr, (same title), *Ann. Geophys. 13*, 583-594, 1995.

Craven, J. D., and L. A. Frank, Latitudinal Motions of the Aurora During Substorms, *J. Geophys. Res. 92*, 4565-4573, 1987.

Crooker, N. U., and A. H. McAllister, Transients Associated with Recurrent Storms, *J. Geophys. Res. 102*, 14041-14047, 1997.

Davies, K., *Ionospheric Radio Propagation*, Dover, New York, 1966.

Dessler, A. J., The Evolution of Arguments Regarding the Existence of Field-Aligned Currents, in *Magnetospheric Currents*, edited by T. A. Potemra, American Geophysical Union, Washington, pp. 22-28, 1983.

de Groot-Hedlin, C. D., and G. Rostoker, Magnetic Signatures of Precursors to Substorm Expansive Phase Onsets, *J. Geophys. Res. 92*, 5845-5856, 1987.

Duffett-Smith, P., *Astronomy with your Personal Computer*, second edition, Cambridge University Press, Cambridge, 258 pp., 1990.

Elphinstone, R. D., J. S. Murphree, D. J. Hearn, W. Heikkila, M. G. Henderson, L. L. Cogger, and I. Sandahl, The Auroral Distribution and its Mapping According to Substorm Phase, *J. Atmos. Terrestr. Phys. 55*, 1741-1762, 1993.

Elphinstone, R. D., D. J. Hearn, L. L. Cogger, J. S. Murphree, H. Singer, V. Sergeev, K. Mursula, D. M. Klumpar, G. D. Reeves, M. Johnson, S. Ohtani, T. A. Potemra, I. Sandahl, E. Nielsen, M. Persson, H. Opgenoorth, P. T. Newell, and Y. I. Feldstein, Observations in the vicinity of substorm onset: Implications for the substorm process, *J. Geophys. Res. 100*, 7937-7969, 1995a.

Elphinstone, R. D., J. S. Murphree, D. J. Hearn, L. L. Cogger, I. Sandahl, P. T. Newell, D. M. Klumpar, S. Ohtani, J. A. Savaud, T. A. Potemra, K. Mursula, A. Wright, and M. Shapshak, The double oval UV auroral distribution 1. Implications for the mapping of auroral arcs, *J. Geophys. Res. 100*, 12075-12092, 1995b.

Elphinstone, R. D., D. J. Hearn, L. L. Cogger, J. S. Murphree, A. Wright, I. Sandahl, S. Ohtani, P. T. Newell, D. M. Klumpar, M. Shapshak, T. A. Potemra, K. Mursula, and J. A. Savaud, The double oval UV auroral distribution 2. The most poleward arc system and the dynamics of the magnetotail, *J. Geophys. Res. 100*, 12093-12102, 1995c.

Fairfield, D. H., Average and Unusual Locations of the Earth's Magnetopause and Bow Shock, *J. Geophys. Res. 76*, 6700-6716, 1971.

Fairfield, D. H., Advances in Magnetospheric Storm and Substorm Research 1989-1991. *preprint*, (Reporter Review Paper presented at the 1991 IAGA Meeting, Vienna, Austria). Laboratory for Extraterrestrial Physics, NASA Goddard Space Flight Center, Greenbelt, November 1991.

Fairfield, D. H., Magnetotail Energy Storage and the Variability of the Magnetotail Current Sheet, in *Magnetic Reconnection in Space and Laboratory Plasmas*, edited by E. W. Hones, Jr., American Geophysical Union, Washington, pp. 168-177, 1984.

Farrugia, C. J., M. P. Freeman, L. F. Burlaga, R. P. Lepping, and K. Takahashi, The Earth's Magnetosphere Under Continued Forcing: Substorm Activity During Passage of an Interplanetary Magnetic Cloud, *J. Geophys. Res. 98*, 7657-7671, 1993.

Forbes, A. J., General Instrumentation, in *Geomagnetism, Volume 1*, edited by J. Jacobs, Academic Press, New York , pp. 51-142, 1987.

Frank, L. A., Dynamics of the Near-Earth Magnetotail - Recent Observations, in *Modeling Magnetospheric Plasma*, edited by T. E. Moore and J. H. Waite, Jr., American Geophysical Union, Washington, pp. 261-276, 1988.

Frank, L. A., C. Y. Huang, and T. E. Eastman, Currents in the Earth's Magnetotail, in *Magnetospheric Currents*, edited by T. A. Potemra, American Geophysical Union, Washington, pp. 147-157, 1983.

Friis-Christensen, E., Polar Cap Current Systems, in *Magnetospheric Currents*, edited by T. A. Potemra, American Geophysical Union, Washington, pp. 86-95, 1983.

Fritz, T. A., and C. A. Neeley, Geostationary Satellites ATS-6 and SMS/GOES: Description, Position, and Data Availability During IMS, in *The IMS Source Book*, edited by C. T. Russell and D. J. Southwood, American Geophysical Union, Washington, pp. 53-64, 1982.

Fujii, R., R. A. Hoffman, P. C. Anderson, J. D. Craven, M. Sugiura, L. A. Frank, and N. C. Maynard, Electrodynamic Parameters in the Nighttime Sector during Auroral Substorms, *J. Geophys. Res. 99*, 6093-6112, 1994.

Fukushima, N., Equivalence in Ground Geomagnetic Effect of Chapman-Vestine's and Birkeland-Alfvén's Electric Current-Systems for Polar Magnetic Storms, *Rep. Ionos. Space Res. Japan 23*, 219-227, 1969.

Fukushima, N., Ground Magnetic Effect of Field-Aligned Currents Connected with Ionospheric Currents, *Geophysics Research Laboratory Report UT-GRL-76-01*, University of Tokyo, 1976.

Gustafsson, G., Corrected Geomagnetic Coordinates for Epoch 1980, in *Magnetospheric Currents*, edited by T. A. Potemra, American Geophysical Union, Washington, pp. 276-283, 1983.

Haines, G. V., Spherical Cap Harmonic Analysis, *J. Geophys. Res. 90*, 2583-2591, 1985.

Hamming, R. W., Numerical Methods for Scientists and Engineers, Second Edition, Dover, New York, 1986.

Hapgood, M. A., Space Physics Coordinate Transformations: A User Guide, *preprint* (later published in *Planetary and Space Science 40*, 711-717, 1992), 17 pp., 1991.

Harrold, B. G. R., Characteristics of Pi 2 Type Geomagnetic Pulsations, *M.Sc. thesis*, University of Alberta, Edmonton, Alberta, Canada, 1986.

Henderson, M. G., and J. S. Murphree, Comparison of Viking Onset Locations with the Predictions of the Thermal Catastrophe Model, *J. Geophys. Res. 100*, 21857-21872, 1995.

Henderson, M. G., G. D. Reeves, R. D. Belian, and J. S. Murphree, Observations of Magnetospheric Substorms Occurring with no Apparent Solar Wind/IMF Trigger, *J. Geophys. Res.101*, 10773-10791, 1996.

Heppner, J. P., and N. C. Maynard, Empirical High-Latitude Electric Field Models, *J. Geophys. Res. 92*, 4467-4489, 1987.

Heppner, J. P., J. D. Stolarik, and E. M. Wescott, Electric Field Measurements and the Identification of Currents Causing Magnetic Disturbances in the Polar Cap, *J. Geophys. Res. 76*, 6028- , 1971.

Hermance, J. F., and G. D. Garland, Deep Electrical Structure under Iceland, *J. Geophys. Res. 73*, 3797-3800, 1968.

Hones, E. W., Jr., C. D. Anger, J. Birn, J. S. Murphree, and L. L. Cogger, A Study of a Magnetospheric Substorm Recorded by the Viking Auroral Imager, *Geophys. Res. Lett. 14*, 411-414, 1987.

Horning, B. L., R. L. McPherron, and D. D. Jackson, Application of Linear Inverse Theory to a Line Current Model of Substorm Current Systems, *J. Geophys. Res. 79*, 5202-5210, 1974.

Hughes, T. J., A Comprehensive Model of Ionospheric-Magnetospheric Current Systems during Periods of Moderate Magnetospheric Activity, , *Ph.D. thesis*, University of Alberta, Edmonton, Alberta, Canada, 1978.

Hughes, T. J., D. W. Oldenburg, and G. Rostoker, Interpretation of Auroral Oval Equivalent Current Flow near Dusk using Inversion Techniques, *J. Geophys. Res. 84*, 450-456, 1979.

Hughes, W. J., The Magnetopause, Magnetotail, and Magnetic Reconnection, in *Introduction to Space Physics*, edited by M. G. Kivelson and C. T. Russell, Cambridge, New York, pp. 91-128, 1995.

Hundhausen, A. J., The Solar Wind, in *Introduction to Space Physics*, edited by M. G. Kivelson and C. T. Russell, Cambridge, New York, pp. 91-128, 1995.

Iijima, T., and T. A. Potemra, Large-Scale Characteristics of Field-Aligned Currents Associated with Substorms, *J. Geophys. Res. 83*, 599-615, 1978.

Iijima, T., M. Watanabe, T. A. Potemra, L. J. Zanetti, and F. J. Rich, The Relationship between Ion and Electron Precipitation Patterns and Field-Aligned Current Systems during a Substorm, in *Magnetospheric Substorms*, edited by J. R. Kan, T. A. Potemra, S. Kokobun, and T. Iijima, American Geophysical Union, Washington, pp. 97-110, 1991.

Jacquey, C., J. A. Savaud, J. Dandouras, and A. Korth, Tailward Propagating Cross-Tail Current Disruption and Dynamics of Near-Earth Tail: A Multi-Point Measurement Analysis, *Geophys. Res. Lett. 20*, 983-986, 1993.

Kamide, Y., Electrodynamic Processes in the Earth's Ionosphere and Magnetosphere, Kyoto Sangyo University Press, Kyoto, 1988.

Kamide, Y., and W. Baumjohann, Estimation of Electric Fields and Currents from International Magnetospheric Study Magnetometer Data for the CDAW 6 Intervals: Implications for Substorm Dynamics, *J. Geophys. Res. 90*, 1305-1317, 1985.

Kamide, Y., and W. Baumjohann, Magnetosphere-Ionosphere Coupling, Springer, Berlin, 1993.

Kamide, Y., and Brekke, Auroral Electrojet Current Density Deduced from the Chatanika Radar and from the Alaska Meridian Chain of Magnetic Observatories, *J. Geophys. Res. 80*, 587, 1975.

Kamide, Y., and S. Kokobun, The Two-Component Auroral Electrojet - The Importance for Substorm Studies, *preprint*, 1996.

Kamide, Y., and C. E. McIlwain, The Onset Time of Magnetospheric Substorms Determined from Ground and Synchronous Satellite Records, *J. Geophys. Res. 79*, 4787-4790, 1974.

Kamide, Y., M. Kanamitsu, and S.-I. Akasofu, A New Method of Mapping Worldwide Potential Contours for Ground Magnetic Perturbations: Equivalent Ionospheric Current Representation, *J. Geophys. Res. 81*, 3810-3820, 1976.

Kawasaki, K., and G. Rostoker, Perturbation Magnetic Fields and Current Systems Associated with Eastward Drifting Auroral Structures, *J. Geophys. Res. 84*, 1464-1478, 1979

King, J. H., Availability of IMP-7 and IMP-8 Data for the IMS Period, in *The IMS Source Book*, edited by C. T. Russell and D. J. Southwood, American Geophysical Union, Washington, pp. 10-20, 1982.

Kisabeth, J. L., The dynamical development of the polar electrojets, *Ph.D. thesis*, University of Alberta, Edmonton, Alberta, Canada, 1972.

Kisabeth, J. L., On Calculating Magnetic and Vector Potential Fields due to Large-Scale Magnetospheric Current Systems and Induced Currents in an Infinitely Conducting Earth, in *Quantitative Modeling of Magnetospheric Processes*, edited by W. P. Olson, American Geophysical Union, Washington, pp. 473-498, 1979.

Kisabeth, J. L., and G. Rostoker, Current Flow in Auroral Loops and Surges Inferred from Ground-Based Magnetic Observations, *J. Geophys. Res. 78*, 5573-5584, 1973.

Kisabeth, J. L., and G. Rostoker, The Expansive Phase of Magnetospheric Substorms 1. Development of the Auroral Electrojets and Auroral Arc Configuration During a Substorm, *J. Geophys. Res. 79*, 972-984, 1974.

Kisabeth, J. L., and G. Rostoker, Modelling of the three-dimensional current systems associated with magnetospheric substorms, *Geophys. J. R. astr. Soc. 49*, 655-683, 1977.

Kivelson, M. G., and C. T. Russell, (editors) *Introduction to Space Physics*, Cambridge, New York, 1995.

Knipp, D. J., Quantifying and Reducing Uncertainy in the Assimilative Mapping of Ionospheric Electrodynamics, *Ph.D. thesis*, University of California, Los Angeles, California, 1989.

Kotikov, A. L., B. D. Bolotinskaya, V. A. Gizler, O. A. Troshichev, O. A. Pashin, and V. R. Tagirov, Structure of Auroral Zone Phenomena from the Data of Meridional Chains of Stations: Magnetic Disturbances in the Nighttime Auroral Zone and Auroras, *J. Atmos. Terr. Physics 53*, 265-274, 1991.

Kotikov, A. L., E. M. Shishkina, O. A. Troshichev, and T. I. Sergienko, Double Structure of Ionospheric Conductivity in the Midnight Auroral Oval during a Substorm, *J. Atmos. Terr. Physics 57*, 177-186, 1995.

Kraus, J. D., and K. R. Carver, *Electromagnetics, 2nd Ed.*, McGraw-Hill, New York, 1973.

Küppers, F., J. Untiedt, W. Baumjohann, K. Lange, and A. G. Jones, A Two-Dimensional Magnetometer Array for Ground-Based Observations of Auroral Zone Electric Currents During the International Magnetospheric Study (IMS), *J. Geophys. 46*, 429-450, 1979.

Lampton, M., Damping-undamping Strategies for the Levenberg-Marquardt Nonlinear Least Squares Method, *Computers in Physics 11*, 110-115, 1997.

Langel, R. A., The Main Field, in *Geomagnetism, Volume 1*, edited by J. Jacobs, Academic Press, New York, pp. 249-512, 1987.

Langel, R. A., T. J. Sabaka, and J. R. Ridgway, *Processing of DMSP Magnetic Data: Handbook of Programs, Tapes and Datasets*, NASA Technical Memorandum 100757, Goddard Spaceflight Center, Greenbelt, 1990.

Lanzerotti, L. J., A. Wolfe, N. Trivedi, C. G. MacLennan, and L. V. Medford, Magnetic Impulse Events at High Latitudes: Magnetopause and Boundary Layer Plasma Processes, *J. Geophys. Res. 95*, 97-107, 1990.

Liou, K., P. T. Newell, C.-I. Meng, and A. T. Y. Lui, Dayside Auroral Activity as a possible Precursor of Substorm Onsets: A Survey using Polar Ultraviolet Imagery, *J. Geophys. Res. 102*, 19835-19843, 1997.

Liu, W. W., and G. Rostoker, On the Origin of Auroral Fingers, *J. Geophys. Res. 98*, 17401-17407, 1993.

Lopez, R. E., H. E. Spence, and C.-I. Meng, DMSP F7 Observations of a Substorm Field-Aligned Current, *J. Geophys. Res. 96*, 19409-19415, 1991.

Lopez, R. E., On the Location of the Substorm Initiation Region, *Adv. Space Res. 13*, 189-198, 1993.

Lopez, R. E., and D. N. Baker, Evidence for Particle Acceleration During Magnetospheric Substorms, *Ap. J. Supp. 90*, 531-539, 1994.

Liu, W. W., and G. Rostoker, Energetic Ring Current Particles Generated by Recurring Substorm Cycles, *J. Geophys. Res. 100*, 21897-21910, 1995.

Lu, G., G. L. Siscoe, A. D. Richmond, T. I. Pulkkinen, N. A. Tsyganenko, H. J. Singer, and B. A. Emery, Mapping of the Ionospheric Field-Aligned Currents to the Equatorial Magnetosphere, *J. Geophys. Res. 102*, 14467-14476, 1997.

Luhmann, J. G., Ionospheres, in *Introduction to Space Physics*, edited by M. G. Kivelson and C. T. Russell, Cambridge, New York, pp. 183-202, 1995.

Lühr, H., H. Geisler, and K. Schlegel, Current Density Models of the Eastward Electrojet Derived from Ground-based Magnetic Field and Radar Measurements, *J. Atmos. Terr. Physics 56*, 81-91, 1994.

Lui, A. T. Y., A Synthesis Model for Magnetospheric Substorms, in *Magnetospheric Substorms*, edited by J. R. Kan, T. A. Potemra, S. Kokobun, and T. Iijima, American Geophysical Union, Washington, pp. 43-60, 1991.

Lundin, R., I Sandahl, J. Woch, and R. Elphinstone, , in *Magnetospheric Substorms*, edited by J. R. Kan, T. A. Potemra, S. Kokobun, and T. Iijima, American Geophysical Union, Washington, pp. 355-373, 1991.

Lyons, L. R., Formation of Auroral Arcs via Magnetosphere-Ionosphere Coupling, *Reviews of Geophysics 30*, 93-112, 1992.

Lyons, L. R., Substorms: Fundamental Observational Features, Distinction from Other Disturbances, and External Triggering, *J. Geophys. Res. 101*, 13011-13025, 1996.

Lyons, L. R., and D. J. Williams, Quantitative Aspects of Magnetospheric Physics, D. Reidel, Dordrecht, 1984.

Marklund, G. T., L. G. Blomberg, T. A. Potemra, J. S. Murphree, F. J. Rich, and K. Stasiewicz, A New Method to Derive "Instantaneous" High-Latitude Potential Distributions from Satellite Measurements Including Auroral Imager Data, *Geophys. Res. Lett. 14*, 439-442, 1987a.

Marklund, G. T., L. G. Blomberg, K. Stasiewicz, J. S. Murphree, R. Pottelette, L. J. Zanetti, T. A. Potemra, D. A. Hardy, and F. J. Rich, Snapshots of High-Latitude Electrodynamics Using Viking and DMSP F7 Observations, *J. Geophys. Res. 93*, 14479-14492, 1987b.

McPherron, R. L., and R. H. Manka, Dynamics of the 1054 UT March 22, 1979 Substorm Event: CDAW 6, *J. Geophys. Res. 90*, 1175-1190, 1985.

McPherron, R. L., Growth Phase of Magnetospheric Substorms, *J. Geophys. Res. 72*, 5592-5599, 1970.

Minatoya, H., N. Sato, T. Saemundsson, and T. Yoshino, Large Displacements of Conjugate Auroras in the Midnight Sector, *J. Geomag. Geoelectr. 48*, 967-975, 1996.

Mishin, V. M., The Magnetogram Inversion Technique and Some Applications, *Space. Sci. Rev. 53*, 83-163, 1990.

Mishin, V. M., L. P. Block, A. D. Bazarzhapov, T. I. Saifudinova, S. B. Lunyushkin, D. Sh. Shirapov, J. Woch, L. Eliasson, G. T. Marklund, L. G. Blomberg, and H. Opgenoorth, A Study of the CDAW 9C Substorm of May 3, 1986, uing Magnetogram Inversion Technique 2, and a Substorm Scenario with Two Active Phases, *J. Geophys. Res. 102*, 19845-19859, 1997.

Murphree, J. S., R. D. Elphinstone, L. L. Cogger, and D. Hearn, Viking Optical Substorm Signatures, in *Magnetospheric Substorms*, edited by J. R. Kan, T. A. Potemra, S. Kokobun, and T. Iijima, American Geophysical Union, Washington, pp. 241-255, 1991.

Murphree, J. S., L. L. Cogger, C. D. Anger, D. D. Wallis, and G. G. Shepherd, Oval Intensifications Associated with Polar Arcs, *Geophys. Res. Lett. 14*, 403-406, 1987.

Nagai, T., Observed Magnetic Substorm Signatures at Synchronous Altitude, *J. Geophys. Res. 87*, 4405-4417, 1982.

National Space Science Data Center, *NSSDC WDC A IM Data Book Supplement 4 (89-17) September 1989*, 1989.

National Space Science Data Center, *NSSDC WDC A IM Data Book Supplement 2 1978-1982 (83-01)*, 1983.

Nishida, A., *Geomagnetic Diagnosis of the Magnetosphere*, Springer, New York, 1978.

Ohtani, S., S. Kokobun, R. C. Elphic, and C. T. Russell, Field-Aligned Current Signatures in the Near-Tail Region 1. ISEE Observations in the Plasma Sheet Boundary Layer, *J. Geophys. Res. 93*, 9709-9720, 1988.

Ohtani, S., K. Takahashi, L. J. Zanetti, T. A. Potemra, and R. W. McEntire, Initial Signatures of Magnetic Field and Energetic Particle Fluxes at Tail Reconfiguration: Explosive Growth Phase, *J. Geophys. Res. 97*, 19311-19324, 1992.

Opgenoorth, H. J., M. A. L. Persson, T. I. Pulkkinen, and R. J. Pellinen, Recovery Phase of Magnetospheric Substorms and its Association with Morning-Sector Aurora, *J. Geophys. Res. 99*, 4115-4129, 1994.

Parkinson, W. D., *Introduction to Geomagnetism*, Elsevier, Amsterdam, 1983.

Parks, G. K., *Physics of Space Plasmas*, Addison-Wesley, Redwood City, California, 1991.

Pangia, M. J., C. S. Lin, and J. N. Barfield, A Correlative Study of Pc 5 Magnetic Pulsations with Substorm Onsets, *J. Geophys. Res. 95*, 10699-10702, 1990.

Persson, M. A. L., H. J. Opgenoorth, and R. D. Elphinstone, On the Difference of Ionospheric Conductances in the Morning and Evening Sectors of the Auroral Oval, *Proceedings of the International Conferecnce on Substorms (ICS-1), Kiruna, Sweden, 23-27 March 1992*, ESA SP-335, 1992.

Potemra, T. A., M. J. Engebretson, L. J. Zanetti, R. E. Erlandson, and B. F. Bythrow, Satellite Observations of Currents and Waves in Space Plasmas, *Laser and Particle Beams 6*, 503-511, 1988.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1988.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Second Edition, Cambridge University Press, Cambridge, 1992.

Priest, E. R., The Sun and its Magnetohydrodynamics, in *Introduction to Space Physics*, edited by M. G. Kivelson and C. T. Russell, Cambridge, New York, pp. 58-90, 1995.

Pulkkinen, T. I., D. N. Baker, N. E. Turner, H. J. Singer, L. A. Frank, J. B. Sigwarth, J. Scudder, R. Anderson, S. Kokobun, R. Nakamura, T. Mukai, J. B. Blake, C. T. Russell, H. Kawano, F. Mozer, and J. A. Slavin, Solar Wind-Magnetosphere Coupling During an Isolated Substorm Event: A Multispacecraft ISTP Study, *Geophys. Res. Lett. 24*, 983-986, 1997.

Pulkkinen, T. I., D. N. Baker, R. J. Pellinen, J. Büchner, H. E. J. Koskinen, R. E. Lopez, R. L. Dyson, and L. A. Frank, Particle Scattering and Current Sheet Stability in the Geomagnetic Tail During the Substorm Growth Phase, *J. Geophys. Res. 97*, 19283-19297, 1992.

Reiff, P. H., G. Lu, J. L. Burch, J. D. Winningham, L. A. Frank, J. D. Craven, W. K. Peterson, and R. A. Heelis, On the High- and Low-Altitude Limits of the Auroral Electric Field Region, in *Auroral Plasma Dynamics*, edited by R. L. Lysak, American Geophysical Union, Washington, pp. 143-154, 1993.

Reeves, G. D., R. D. Belian, and T. A. Fritz, Numerical Tracing of Energetic Particle Drifts in a Model Magnetosphere, *J. Geophys. Res. 96*, 13997-14008, 1991.

Rice, J. R., *Numerical Methods, Software, and Analysis*, McGraw-Hill, New York, 1983.

Rich, F. J., D. A. Hardy, R. H. Redus, and M. S. Gussenhoven, Northward IMF and Patterns of High-Latitude Precipitation and Field-Aligned Currents: The February 1986 Storm, *J. Geophys. Res. 95*, 7893-7913, 1990.

Richmond, A. D., Assimilative Mapping of Ionospheric Electrodynamics, *Adv. Space Res. 12*, 59-68, 1992.

Richmond, A. D., and W. Baumjohann, Three-dimensional Analysis of Magnetometer Array Data, *J. Geophys. 54*, 138-156, 1983.

Richmond, A. D., and Y. Kamide, Mapping Electrodynamic Features of the High-Latitude Ionosphere from Localized Observations: Technique, *J. Geophys. Res. 93*, 5471-5759, 1988.

Richmond, A. D., Y. Kamide, S.-I. Akasofu, D. Alcaydé, M. Blanc, O. de la Beaujardière, D. S. Evans, J. C. Foster, E. Friis-Christensen, J. M. Holt, R. J. Pellinen, C. Senior, and A. N. Zaitzev, Global Measures of Ionospheric Electrodynamic Activity Inferred from Combined Incoherent Scatter Radar and Ground Magnetometer Observations, *J. Geophys. Res. 95*, 1061-1071, 1990.

Ridgway, J. R., T. J. Sabaka, D. Chinn, and R. A. Langel, Processing of DMSP Magnetic Data and its Use in Geomagnetic Field Modeling, NASA Technical Memorandum 100750, NASA Goddard Space Flight Center, Greenbelt, 1989.

Rostoker, G., Midlatitude Transition Bays and their Relation to the Spatial Movement of Overhead Current Systems, *J. Geophys. Res. 71*, 79-95, 1966.

Rostoker, G., Classification of Polar Magnetic Disturbances, *J. Geophys. Res. 74*, 5161-5168, 1969.

Rostoker, G., Current Flow in the Magnetosphere During Magnetospheric Substorms, *J. Geophys. Res. 79*, 1994-1998, 1974.

Rostoker, G., Magnetospheric and Ionospheric Currents in the Polar Cusp and their Dependence on the B_y Component of the Interplanetary Magnetic Field, *J. Geophys. Res. 85*, 4167-4176, 1980.

Rostoker, G., Magnetospheric Substorms - Their Phenomenology and Predictability, in *Solar-Terrestrial Predictions - IV*, Volume 1, edited by J. Hruska, M. A. Shea, D. F. Smart, and G. Heckman, NOAA, Boulder, pp. 21-35, 1993.

Rostoker, G., The Phenomenology and Physics of Magnetospheric Substorms, *J. Geophys. Res. 101*, 12955-12973, 1996.

Rostoker, G., On the Place of the Pseudo-Breakup in a Magnetospheric Substorm, *Geophys. Res. Lett., in press*, 1998.

Rostoker, G., and T. D. Phan, Variation of Auroral Electrojet Spatial Location as a Function of the level of Magnetospheric Activity, *J. Geophys. Res. 91*, 1716-1722, 1986.

Rostoker, G., C. W. Anderson III, D. W. Oldenburg, P. A. Camfield, and D. I. Gough, Development of a Polar Magnetic Substorm Current System, *J. Geophys. Res. 75*, 6318-6323, 1970.

Rostoker, G., J. C. Samson, F. Creutzberg, T. J. Hughes, D. R. McDiarmid, A. G. McNamara, A. Vallance Jones, D. D. Wallis, and L. L. Cogger, CANOPUS - A Ground-Based Instrument Array for Remote Sensing the High Latitude Ionosphere During the ISTP/GGS Program, *Space Sci. Rev. 71*, 743-760, 1995.

Rostoker, G., I. Spadinger, and J. C. Samson, Local Time Variation in the Response of Pc 5 Pulsations in the Morning Sector to Substorm Expansive Phase Onsets near Midnight, *J. Geophys. Res. 89*, 6749-6757, 1984.

Rostoker, G., A. Vallance-Jones, R. L. Gattinger, C. D. Anger, and J. S. Murphree, The Development of the Substorm Expansive Phase: The "Eye" of the Substorm, *Geophys. Res. Lett. 14*, 399-402, 1987.

Russell, C. T., Geophysical Coordinate Transformations, *Cosmic Electrodynamics 2*, 184-196, 1971.

Russell, C. T., N. U. Crooker, H. C. Koons, G. Rostoker, D. Sibeck, E. Szuszczewicz, M. Teague, and V. Papitashvili, Magnetospheric Working Group Report Solar-Terrestrial Predictions Workshop, Ottawa, May 1992, in *Solar-Terrestrial Predictions - IV*, Volume 1, edited by J. Hruska, M. A. Shea, D. F. Smart, and G. Heckman, NOAA, Boulder, pp. 49-58, 1993.

Rycroft, M. J., Some Aspects of Geomagnetically Conjugate Phenomena, *Ann. Geophys. 5*, 463-478, 1987.

Sabatier, P. C., The Inverse Problem, *Public Lecture*, Department of Physics, University of Alberta, September 27, 1993.

Samson, J. C., Geomagnetic Pulsations and Plasma Waves in the Earth's Magnetosphere, in *Geomagnetism, Volume 4*, edited by J. Jacobs, Academic Press, New York , pp. 481-592, 1991.

Samson, J. C., and G. Rostoker, Response of Dayside Pc 5 Pulsations to Substorm Activity in the Nighttime Magnetosphere, *J. Geophys. Res. 86*, 733-752, 1981.

Samson, J. C., D. D. Wallis, T. J. Hughes, F. Creutzberg, J. M. Ruohoiemi, and R. A. Greenwald, Substorm Intensifications and Field Line Resonances in the Nightside Magnetosphere, *J. Geophys. Res. 97*, 8495-8518, 1992.

Sato, M., Y. Kamide, A. D. Richmond, A. Brekke, and S. Nozawa, Regional Estimation of Electric Fields and Currents in the Polar Ionosphere, *Geophys. Res. Lett. 22*, 283-286, 1995.

Savaud, J.-A., and J. R. Winckler, Dynamics of Plasma, Energetic Particles, and Fields Near Synchronous Orbit in the Nighttime Sector During Magnetospheric Substorms, *J. Geophys. Res. 85*, 2043-2056, 1980.

Senior, C., R. M. Robinson, and T. A. Potemra, Relationship Between Field-Aligned Currents, Diffuse Auroral Precipitation and the Westward Electrojet in the Early Morning Sector, *J. Geophys. Res. 87*, 10469-10477, 1982.

Sergeev, V. A., L. I. Vagina, R. D. Elphinstone, J. S. Murphree, D. J. Hearn. L. L. Cogger. and M. L. Johnson, Comparison of UV Optical Signatures with the Substorm Current Wedge as Predicted by an Inversion Algorithm, *J. Geophys. Res. 101*, 2615-2627. 1994

Skone, S., Characterizing the Magnetic Field at Geostationary Orbit, *M.Sc. thesis*, University of Alberta, Edmonton, Alberta, Canada, 1994.

Stern, D. P., The Beginning of Substorm Research, in *Magnetospheric Substorms*, edited by J. R. Kan, T. A. Potemra, S. Kokobun, and T. Iijima, American Geophysical Union. Washington, pp. 11-14, 1991.

Stüdemann, W., B. Wilken, G. Kremser, A. Korth, J. F. Fennell, B. Blake, R. Koga, D. Hall, D. Bryant, F. Søraas, K. Brønstad, T. A. Fritz, R. Lundin, and G. Gloeckler, The May 2-3, 1986 Magnetic Storm: First Energetic Ion Composition Observations with the MICS Instrument on Viking, *Geophys. Res. Lett. 14*, 455-458, 1987.

Sofko, G. J., A. Korth, and G. Kremser, The August 28, 1978, Storm 1. GEOS 2 Observations of the Initial Magnetopause Crossings and STARE Observations Near the Ionospheric Convection Reversal, *J. Geophys. Res. 90*, 1449-1457, 1985.

Sulzbacher, H., W. Baumjohann, and T. A. Potemra, Coordinated Magnetic Observations of Morning Sector Auroral Zone Currents with Triad and the Scandinavian Magnetometer Array: A Case Study, *J. Geophys. 48*, 7-17, 1980.

Sulzbacher, H., W. Baumjohann, T. A. Potemra, E. Nielsen, and G. Gustafsson, Observations of Ionospheric and Field-Aligned Currents in the Late Afternoon Sector with Triad and the Scandinavian Magnetometer Array, *J. Geophys. 51*, 55-65, 1982.

Sun, W., B.-H. Ahn, S.-I. Akasofu, and Y. Kamide, A Comparision of the Observed Mid-Latitude Magentic Disturbance Fields with those Reproduced from the High-Latitude Modeling Current Systems, *J. Geophys. Res. 89*, 10881-10889, 1984.

Temple, R., The Genius of China, Simon and Schuster, New York, 1986.

Tsyganenko, N. A., An Empirical Model of the Substorm Current Wedge, *J. Geophys. Res. 102*, 19935-19941, 1997.

Untiedt, J., and W. Baumjohann, Studies of Polar Current Systems using the IMS Scandinavian Magnetometer Array, *Space Sci. Rev. 63*, 245-390, 1993.

U. S. Navy Hydrographic Office, Magnetic Variation Epoch 1965, H. O. 1706-N (map), Washington, May 1965.

Vacquier, V., Many Jobs, *Ann. Rev. Earth and Plan. Sci. 21*, 1-18, 1993.

Vallance Jones, A., R. L. Gattinger, F. Creutzberg, R. A. King, P. Prikryl, L. L. Cogger, D. J McEwen, F. R. Harris, C. D. Anger, J. S. Murphree, and R. A. Koehler, A Comparison of CANOPUS Ground Optical Data with Images from the Viking UV Camera, *Geophys. Res. Lett. 14*, 391-394, 1987.

Van Zandt, T. E., W. L. Clark, and J. M. Warnock, Magnetic Apex Coordinates: A Magnetic Coordinate System for the Ionospheric F_2 Layer, *J. Geophys. Res. 77*, 2406-2411, 1972.

Vasyliunas, V. M., Mathematical Models of Magnetospheric Convection and its Coupling to the Ionosphere, in *Particles and Fields in the Magnetosphere*, edited by B. M. McCormac, Springer, New York, pp. 60-71, 1970.

Walker, J. K., V. Yu. Semenov, and T. L. Hansen, Synoptic Models of High Latitude Magnetic Activity and Equivalent Ionospheric and Induced Currents, *J. Atmos. Terrest. Phys. 59*, 1435-1452, 1997.

Walker, R. J., and C. T. Russell, Solar Wind Interactions with Magnetized Planets, in *Introduction to Space Physics*, edited by M. G. Kivelson and C. T. Russell, Cambridge, New York, pp. 164-182, 1995.

Wiens, R. G., and G. Rostoker, Characteristics of the Development of the Westward Electrojet During the Expansive Phase of Magnetospheric Substorms, *J. Geophys. Res. 80*, 2109-2128, 1975.

Wilkinson, D., and G. Ushomirskiy, GOES Space Environment Monitor CD-ROM User Documentation, National Geophysical Data Center, Boulder, 1994.

Wolf, R. A., Magnetospheric Configuration, in *Introduction to Space Physics*, edited by M. G Kivelson and C. T. Russell, Cambridge, New York, pp. 288-329, 1995.

Wolfram, S., Mathematica: A System for doing Mathematics by Computer, Second Edition, Addison-Wesley, Redwood City, 1991.

Xu, B.-L., J. C. Samson, W. W. Liu, F. Creutzberg, and T. J. Hughes, Observations of Optical Aurora Modulated by Resonant Alfvén Waves, *Report 93-06*, Canadian Network for Space Research, University of Alberta, Edmonton, Alberta, Canada, 1993.

Yabuzaki, T., and T. Ogawa, Rocket Measurement of Sq Ionospheric Currents over Kagoshima, Japan, *J. Geophys. Res. 79*, 1999-2001, 1974.

Yeoman, T. K., D. K Milling, and D. Orr, Pi 2 Pulsational Polarization Patterns on the U. K. Sub-Auroral Magnetometer Network (SAMNET), *Planet. Space Sci. 38*, 589-602, 1990.

Zanetti, L. J., W. Baumjohann, and T. A. Potemra, Ionospheric and Birkeland Current Distributions Inferred from the MAGSAT Magnetometer Data, *J. Geophys. Res. 88*, 4875 4884, 1983.

Zmuda, A. J., J. H. Martin, and F. T. Heuring, Transverse Magnetic Disturbances at 1100 Kilometers in the Auroral Region, *J. Geophys. Res. 71*, 5033-5045, 1966.

Zmuda, A. J., J. H. Martin, and F. T. Heuring, Dayside Magnetic Disturbances at 1100 Kilometers in the Auroral Oval, *J. Geophys. Res. 72*, 1115-1117, 1967.